

[illegible]



Genome version 4.5  
Copyright (c) 1994 - 2000 Genomics Ltd

CM protein - protein search, using sw model

Run on: August 31, 2001, 15:00:32; Search time 59.12 seconds

(without alignments)  
796,695 Million cell updates/sec

Title: US-09-577-657A-1  
Perfect score: 1847  
Sequence: 1 FMNRGESSYAONSSFTQ2 LEAKIPKTTIILVSLK 356

Scoring table:  
R10SUM62  
Gapop 10 0, Gapext 0 5

Searched: 425026 seqs, 12339527 residues

Total number of hits satisfying chosen parameters: 425026

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database:

SPRENBL\_16:  
1: sp.archaea:  
2: sp.bacteria:  
3: sp.fungi:  
4: sp.human:  
5: sp.invertebrate:  
6: sp.mammal:  
7: sp.mhc:  
8: sp.organelle:  
9: sp.phage:  
10: sp.plant:  
11: sp.prokaryote:  
12: sp.unclassified:  
13: sp.vertebrate:  
14: sp.virus:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1847	100.0	369	10 Q9FZN8	Q9FZN8 camellia sl
2	724	39.2	392	10 Q9SBK6	Q9SBK6 brassica ra
3	713	38.6	359	10 Q9SPV4	Q9SPV4 clarkia dre
4	637.5	34.5	354	10 Q9FJ26	Q9FJ26 arabidopsis
5	630	34.1	371	10 Q9J234	Q9J234 arabidopsis
6	621.5	33.6	364	10 Q9FJ29	Q9FJ29 arabidopsis
7	585.5	31.7	363	10 Q9LS20	Q9LS20 arabidopsis
8	582.5	31.5	368	10 Q9LPL5	Q9LPL5 arabidopsis
9	570	30.9	359	10 Q9ZPT3	Q9ZPT3 arabidopsis
10	499.5	27.0	351	10 Q9X157	Q9X157 arabidopsis
11	496.5	26.9	348	10 Q9FYC4	Q9FYC4 arabidopsis
12	485.5	26.3	348	10 Q9FYC6	Q9FYC6 arabidopsis
13	482.5	26.1	359	10 Q9LS10	Q9LS10 arabidopsis
14	482	26.0	374	10 Q9FKC8	Q9FKC8 arabidopsis
15	481	26.0	386	10 Q9ELN8	Q9ELN8 arabidopsis
16	480	26.0	379	10 Q9FYC3	Q9FYC3 arabidopsis
17	480	25.0	319	10 Q9F552	Q9F552 arabidopsis
18	475.5	25.7	323	10 Q9FWR4	Q9FWR4 arabidopsis
19	475.5	25.7	361	10 Q9FKK0	Q9FKK0 arabidopsis

20	466.5	25.3	380	10 Q9FWJ1	Q9FWJ1 oryza sativ
21	454	24.6	362	10 Q9FKU0	Q9FKU0 arabidopsis
22	436.5	24.6	380	10 Q9FWJ2	Q9FWJ2 arabidopsis
23	424	23.0	353	10 Q9M4F2	Q9M4F2 brassica na
24	387	21.0	318	10 Q9FT13	Q9FT13 brassica na
25	105.5	5.7	100	10 Q9LS11	Q9LS11 arabidopsis
26	101.5	5.5	1016	3 Q9UT45	Q9UT45 schizosacch
27	99	5.4	319	5 Q9Z637	Q9Z637 caenorhabdi
28	98	5.3	1236	5 Q44875	Q44875 caenorhabdi
29	97	5.3	1494	5 Q21436	Q21436 caenorhabdi
30	96	5.2	1280	2 Q9ZAJ5	Q9ZAJ5 clostridium
31	96	5.2	1862	10 Q9SHM5	Q9SHM5 arabidopsis
32	94.5	5.1	469	14 Q9WBS2	Q9WBS2 physalis se
33	94.5	5.1	590	2 Q54535	Q54535 streptococ
34	94.5	5.1	787	1 Q58025	Q58025 pyrococcus
35	94.5	5.1	3325	14 Q9IFR9	Q9IFR9 turkey herp
36	94.5	5.1	3442	14 Q56N33	Q56N33 turkey herp
37	94	5.1	983	10 Q9ZUK3	Q9ZUK3 arabidopsis
38	93.5	5.1	525	1 Q56N40	Q56N40 schizosacch
39	93	5.0	295	2 Q9XDK6	Q9XDK6 bacteroides
40	93	5.0	736	5 Q25856	Q25856 plasmodium
41	93	5.0	749	5 Q02602	Q02602 plasmodium
42	93	5.0	897	10 Q9LHC0	Q9LHC0 arabidopsis
43	92.5	5.0	2185	14 Q9E7C3	Q9E7C3 human coxa
44	92.5	5.0	2185	14 Q9E7C2	Q9E7C2 human coxa
45	92	5.0	392	3 Q94240	Q94240 schizosacch

## ALIGNMENTS

RESULT	1	PRELIMINARY:	PRT:	369 AA.
Q9FZN8	Q9FZN8			
AC	Q9FZN8			
DT	01-MAR-2001 (TRENBLER, 16, Created)			
DT	01-MAR-2001 (TRENBLER, 16, Last sequence update)			
DT	01-MAR-2001 (TRENBLER, 16, Last annotation update)			
DE	CAFFEINE SYNTHASE.			
GN	TCSL			
OS	Camellia sinensis (Tea).			
OC	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;			
OC	Magnoliophyta; eudicotyledons; core eudicots; Asteridae; Ericales;			
OC	Theaceae; Camellia.			
OX	NCBI_TaxID=4442;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
FX	MEDLINE=2043735; PubMed=10984041;			
FA	Kato M, Mizuno K, Crozier A, Fujimura T, Ashihara H.;			
RT	"Caffeine synthase gene from tea leaves.";			
RL	Nature 406:956-957(2000).			
DR	EMBL: AB031280; BMB12278.1;			
SO	SEQUENCE 369 AA; 417.7 MW; E6D26.087E47576 CRC64;			
Query Match	100.0%; Score 1847; DB 10; Length 369;			
Best Local Similarity	100.0%; Pred. No. 1.1e-144;			
Matches 350; Conservative	U. Mismatches	U. Indels	0. Gaps	0;
QY	1 FMNRGESSYAONSSFTQVASMALPNAVEFLFSRDFHLQALNADGCAAGPNTF 60			
DB	14 FMNRGESSYAONSSFTQVASMALPNAVEFLFSRDFHLQALNADGCAAGPNTF 73			
QY	51 AVISTIKRPMKKCPRLNGLTLEIYVINDJFGNPNFLFKGLTSEVYGNCEPVCVM 120			
DB	74 AVISTIKRPMKKCPRLNGLTLEIYVINDJFGNPNFLFKGLTSEVYGNCEPVCVM 133			
QY	121 GVGSEHGRFLFPNSLHLHSSYSVHMLTQAKGILTSEGLALNKGITYIKTSPVPE 180			
DB	134 GVGSEHGRFLFPNSLHLHSSYSVHMLTQAKGILTSEGLALNKGITYIKTSPVPE 193			
QY	141 AVISQPHFNTMTLAPSCFVVPN3CMVITLPDQCSFPEMCSFTWELLAMATAETVS 240			

Db 194 ATLSQFHEDETFMFLNARSOEVVPCNGWVLLRGQCSPPDSMOGCTWELLAMATAELVS 253  
 QY 241 OGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGFDLDSVEMQENDKWVPSGK 300  
 Db 254 OGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGFDLDSVEMQENDKWVPSGK 313  
 QY 301 FTKVRAATEPTISQGFELMDKLYDKFTHVVSOLFAPKPTTSTILVLSKIDG 356  
 Db 314 FTKVRAATEPTISQGFELMDKLYDKFTHVVSOLFAPKPTTSTILVLSKIDG 369

RESULT 2  
 Q9SRK6 PRELIMINARY: PRT: 392 AA.  
 ID O9SRK6  
 AC O9SRK6  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 16, Last annotation update)  
 DE FLORAL NECTARY-SPECIFIC PROTEIN.  
 OS Brassica rapa subsp. pekinensis (Chinese cabbage) (Celery cabbage).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Brassica.  
 OX NCBI\_TaxID=51351;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Song J.T., Seo H.S., Song S.L., Lee J.S., Choi Y.D.;  
 RT "Characterization of a novel gene expressed specifically in the floral  
 nectaries of Brassica campestris L. ssp. pekinensis.";  
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.  
 DF EMBL: AF179222; AAF22389.1; ...  
 SQ SEQUENCE 392 AA, 43815 MW, 253785305365757 CRC64;

Query Match 39.2%; Score 724; DB 10; Length 392;  
 Best local Similarity 38.1%; Pred. No. 7.9e-52;  
 Matches 146; Conservative 78; Mismatches 129; Indels 30; Gaps 5;

QY 2 MNRGESSYAONSFTQOVASMAQPALENAVETIFSRDHFLOALNAADLGCAGPNTFA 61  
 Db 9 MNKGNGETSTAKNSIVOSNITSIGRVMDEALKLMIRNSELISFGIADLGSSGPNLSL 68  
 QY 62 VISTIKRMKCKKELNCJULEVLYLNDLFGNLFNLGKLS-----EVINKKCEV- 115  
 Db 69 SISNIVETIONLCCHDRPVPBELSLNDLPSDENPITFASLEPFDRVKKRNNNESLG 128  
 QY 116 -----PCYVMGVPGSFHGRILFPNRLSLVHSSYSYHMLTQAPKG-LISREGLAL---- 163  
 Db 129 FEHSGGPGCFVSAVPSGFYGRFLPRRLSLHFVHSSSLHMLSOVPGCEVNNKDKGVITADL 188  
 QY 164 -NMGKTIYSTSPVYVEALYSOFHEDETFMFLNARSOEVVPCNGWVLLRGQCSPPDSM 222  
 Db 189 DNGKTIYSTSKTSKSAHKAHYVALOPDTDFSYFLSPSEFELVPGSRMVLSPGLRSSPPTTE 248  
 QY 223 QSCFTWELLAMATAELVSOGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGF 282  
 Db 249 ESCYOWELLAOALMSLAKEGIIIEENIDAFNADYAAAPSELMALIEKESPSIDLLEIS 308  
 QY 283 DLD-----SVEMQENDKWVGEKFTKYVRAATEPTISNOGPELMKLYDKFTH 331  
 Db 309 PYVMGEGSSISDSDSYDVRKPRPEALASGRVAKTIRAVVEPMLEPTFGORVMDLEFERVAK 368  
 QY 332 IVVSOLFAPKPTTSTILVLSKIDG 354  
 Db 369 LVGEYVYVSSPRTIYIVLSLRM 391

RESULT 3  
 Q9SPV4 PRELIMINARY: PRT: 359 AA.  
 ID O9SPV4  
 AC O9SPV4  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
 DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE.  
 GN SMT.  
 OS Clarkia breweri.  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Myrtales; Onagraceae; Clarkia.  
 OX NCBI\_TaxID=36903;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Ross J.F., Nam K.H., D'Auria J., Pichersky E.;  
 RT "S-Adenosyl-L-methionine:salicylic acid carboxyl methyltransferase, an  
 enzyme involved in floral scent production and plant defense,  
 represents a new class of plant methyltransferases.";  
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 DF EMBL: AF133053; AAF00108.1; ...  
 KM Transferase; Methyltransferase.  
 SQ SEQUENCE 359 AA, 40289 MW, BOFE3E41AE8EDB51 CRC64;

Query Match 38.6%; Score 713; DB 10; Length 359;  
 Best local Similarity 41.5%; Pred. No. 5.7e-51;  
 Matches 152; Conservative 68; Mismatches 110; Indels 36; Gaps 8;

QY 2 MNRGESSYAONSFTQOVASMAQPALENAVETIFSRDHFLOALNAADLGCAGPNTFA 60  
 Db 9 MNKGAGENSYANMSFIQROVISTIKRTEAATITLISGVTTRIALADLGSSGNALF 68  
 QY 64 AV-----ISTIKRMKCKKELNCJULEVLYLNDLFGNDFNLFLKLSSEVIQNKCEVPC 117  
 Db 69 ATELTKTYEELRKMKGR-----NSPEYDFELNDLPGNFALFISLPIF---NDVPGV-C 121  
 QY 118 YVMGVPGSFHGRILFPNRLSLVHSSYSYHMLTQAPKGLTSRGLALNKGKIYSTSPV 177  
 Db 122 FINGVGSFYGLFLPNTLTHFIHSSYSMLMSOVIGIES-----NKGNTIMANTQPOS 175  
 QY 178 VDEATYSOFHEDETFMFLNARSOEVVPCNGWVLLRGQCSPPDSMOGCTWELLAMATAEL 237  
 Db 176 VLNAYYKQFOEDHALFLRRAQEVVPGRMVLTILGPRSEDRASTRECLLIWOLLAMALNQ 235  
 QY 238 IYSOGLIDEDKLDTPNIPSYASLEEVKDIYERDGSFTIDHIEGDL-----D 285  
 Db 236 MNSGEIIEKKDKKNINQYTSPTEVAEILKESGFLIDHASEIYSSCTKDGSGGG 295  
 QY 286 SVEMQENDKWVGEKFTKYVRAATEPTISNOGPELMKLYDKFTHIVVSOLFAPKPTT 345  
 Db 296 SVEEE-----GYVVARCMRAVAEPILLDFHGEALIEDVFRYKLLIEMSKETKFI 348  
 QY 346 STILVL 351  
 Db 349 NVIVSL 354

RESULT 4  
 Q9FJZ6 PRELIMINARY: PRT: 354 AA.  
 ID Q9FJZ6  
 AC Q9FJZ6  
 DT 01-MAY-2001 (TREMBlrel. 16, Created)  
 DT 01-MAY-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAY-2001 (TREMBlrel. 16, Last annotation update)  
 DE S-ADENOSYL-L-METHIONININ:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE-LIKE  
 PROTEIN.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;  
 OC Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-COLUMBIA;  
 MEDLINE=98403884; PubMed=9734815;  
 Kotani H., Nakamura Y., Sato S., Asamizu E., Kaneko T., Miyajima N.,





QY 122 KISF FVVRVAVLSQFHEDFTMLNAR-----SQGVVPCMGV-----LLIKKQVSD 218  
 DB 96 KTSYTLTAELTAEDVEFKANSSLLNALLSTINERLYHNDKLEKVLISFGASNELE 155  
 QY 219 PSMQSCFTWELMAMAEIAYS QGLIDEKIDFENIPSYFASLEFKDIVERGSDTD 277  
 DB 156 ENPIIAPYDRLFPKVPVGIJSCENLVKLIKDEPKPKTTISIKELPKMGEKANVDIE 215  
 QY 278 HIEHDLQ-SVEKQENKRWGCKFTKVRVAF-TEPIISQSGPEIMDKLYDKFTHIYVS 335  
 DB 216 NIEVLVYIKKKLSQNHVYISPEPKSKVAKIKCFAYVINGPPEATED--LETLPIIFWD 273  
 QY 446 DLEAKIKFTTSILVLSKI 354  
 DB 274 DIDD-----ILIVSKV 284

RESULT 2  
 KHL2\_RIBOSA  
 ID RHL2\_RIBOSA STANDARD; PRT: 458 AA.  
 AC P50922;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DB RIBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN 2 (EC 4.1.1.39) (RUBISCO LARGE SUBUNIT)  
 DB CMBL2 OR CRBM OR PPPI  
 DB Bacteroid capsulatus (Rhodospirillum rubrum)  
 DB Bacteroid proteobacterium; alpha subdivision; Rhodobacter group;  
 DB Rhodobacter  
 DB MBL\_TaxID=10611;  
 DB 11

SEQUENCE FROM N.A.  
 STRAIN=ATCC 11166;  
 RA Lattimer F.W., Lu T.-Y.S., Buley D.M.;  
 RT "Sequence and expression of the form II ribulose-bisphosphate  
 KT carboxylase/oxygenase (RUBISCO) gene from Rhodobacter capsulatus";  
 PE FASEB J. 9: A1275-A1275(1997).

CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
 D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
 CC ACTIVE SITE.

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =  
 CC 2-3-PHOSPHO-D-GLYCERATE.

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =  
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.

CC -1- SUBUNIT: IN CONTRAST TO FORM I GENES, THE FORM II RUBISCO ARE  
 CC COMPOSED OF ONLY LARGE SUBUNITS (BY SIMILARITY).

CC -1- MISCELLANEOUS: THIS PROTEIN IS ENCODED WITHIN THE FORM II  
 CC RIBULOSE-BISPHOSPHATE CARBOXYLASE OPERON.

CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL consortium -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announcements/)  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).

DB EMBL: U21445; AAR92348.1;  
 DB HSSP: P04718; 2R0S;  
 DB InterPro: IPR000685;  
 DB Pfam: PF00016; RUBISCO\_Large; 1;  
 DB PROSITE: PS00167; RUBISCO\_Large; 1;  
 KW Photosynthesis; Carbon dioxide fixation; Photorespiration;  
 KW Lysate; Oxidoreductase; Monooxygenase; Multiene family;  
 FT ACT\_SITE 191 191  
 FT BINDING OF CO(2) ACTIVATES THE ENZYME  
 FT (BY SIMILARITY).  
 SO SOURCE 458 AA; 50091 MW; 26687057653FAA Prev 4.

Query Match 5.38; Score 97.5; DB 1; Length 458;  
 Best Local Similarity 19.2%; Pred. No. 1.4;  
 Matches 70; Conservative 59; Mismatches 147; Indels 59; Gaps 17;

QY 3 NQGHSSVYALNSST-----QGVASM-----AGVALN- AVEPLPSPDRL 44  
 DB 112 NQGHSSVYALNSST-----QGVASM-----AGVALN- AVEPLPSPDRL 44  
 QY 44 QALNADLC-----AASPNTFAVISTIKPMERKQELNLTLEQVYIN 89  
 DB 172 PPKPEAD-ACVEHPLQSDPTKNIHQGNQJFATLEKILPLVADAMRAQVHGEAKLISA 240  
 QY 90 DIFGND-FNLIPLG-LSSEVQDKRTEVYVYVW-VYQ-----SPRPIPLPNSLTHES 142  
 DB 231 NITAEHVEVAVAGVYLELISGNAHVALVYQVYVGAALIIAKSKPQGLHYHAG 290  
 QY 143 YSVHMLTQAFKLTSPGELALNGCYVTEKTPPVVPEVAVTSQFHEDFTMLNARSGV 202  
 DB 291 HGATSPQSMKQVTA-----EVLISMS-----HIGASIHIGMCT--GEMEDIA 344  
 QY 203 FNSVMVILHQPQVSNGLMSFTWELMAMAEIAYSQGLIDEKIDFENIPSYFASLE 252  
 DB 445 STKIMAVMLNDEAAGDPFYHQD--W-LSKATLPIISG-----MNALPLQCFDNIAG 384  
 QY 263 EVKQIVEPEFSTLHGFPHLISVEMQEN IAWVPCIKFYVVAPEPIPLISQAFGEI 421  
 DB 385 H-SVVTQSGAGPGLDGTAGAKSLPQSCAMKAGVILVYAKSHRILANAFESFND 443  
 QY 322 MDKLY 326  
 DB 444 ADKLY 448

RESULT 3  
 POLG\_ECOL2T  
 ID POLG\_ECOL2T STANDARD; PRT: 2193 AA.  
 AC Q66575; Q66576;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)

DB GENOME POLYPROTEIN (CONTAINS: COAT PROTEIN VP4 (P1A); COAT PROTEIN VP2  
 DE (P1B); COAT PROTEIN VP3 (P1C); COAT PROTEIN VP1 (P1D); POLYOMAIN 2A  
 DE (P2); COAT PROTEIN P2B; COAT PROTEIN P2C; COAT PROTEIN P2D; COAT PROTEIN  
 DE (P2E); GENOME-LINKED PROTEIN VPG (P3B); POLYOMAIN 3C (P3); POLYOMAIN 3D  
 DE (P3E); POLYOMAIN 3F (P3G); RNA-DIRECTED RNA POLYMERASE (P4.7.7.48) (P4D);  
 OS Echovirus 12 (Strain Travis).

CC Viruses; ssRNA positive strand viruses, no DNA stage; Picornaviridae;  
 CC Enterovirus.

CC NCBI\_TaxID=103909;  
 DB 11

SEQUENCE FROM N.A.  
 STRAIN=WILDTYPE;  
 RX MEDLINE=95364006; PubMed=7637932;

FA Kraus W., Zimmermann H., Zimmermann A., Lehrs H.J., Nelson-Salz B.;  
 PT Infectious cDNA clones of echovirus 12 and a variant resistant  
 RT against the uncoding inhibitor rhodamine differ in seven amino  
 RT acids.\*

PL 1. Vitro 64-5853-5858(1995).

CC -1- FUNCTION: P2A AND THE P2C POLYPEPTIDES ARE PROTEASES THAT CLEAVE  
 CC AT CERTAIN Q/G SITES IN THE POLYPROTEIN. THEY ARE CYSTEINE  
 CC PROTEASES.

CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 PROSOMERICAL UNITS,  
 CC EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2,  
 CC VP3, AND VP4.

CC -1- PTM: SPECIFIC ENZYMOLOGICAL CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC CLEAVAGE BETWEEN VP4 AND VP2 IS CATALYZED BY VP2. CLEAVAGE BETWEEN  
 CC VP2 AND VP3 IS CATALYZED BY VP3. ALL OTHER CLEAVAGES ARE CATALYZED BY P2C.

CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -1- SIMILARITY: P2C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL consortium -

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in the public domain. This statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.embnet.org/licenses/> or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL: X79647: CAAS5650.1  
EMBL: X77708: CAAS4783.1  
MEROPS: C03.001  
MEROPS: C03.020  
InterPro: IPR000681  
InterPro: IPR000199  
InterPro: IPR000605  
InterPro: IPR001205  
InterPro: IPR001643  
InterPro: IPR002527  
Pfam: PF00548: Cys-Protease\_3c\_1  
Pfam: PF00947: Pico\_P2B\_1  
Pfam: PF01552: Pico\_P2B\_1  
Pfam: PF00680: RNA\_drp\_RNA\_pol\_1  
Pfam: PF00910: RNA\_helicase\_1  
Pfam: PF00073: Ruv\_3  
PRINTS: PR06194: CALF\_VIF\_PNS  
Polyprotein: Coat protein: Core protein: Transferase: Myristate;  
RNA-directed RNA polymerase: Hydrolase: Thiol protease;  
CHAIN 2  
CHAIN 360  
CHAIN 331  
CHAIN 568  
CHAIN 569  
CHAIN 860  
CHAIN 1010  
CHAIN 1108  
CHAIN 1437  
CHAIN 1438  
CHAIN 1526  
CHAIN 1527  
CHAIN 1549  
CHAIN 1731  
CHAIN 1732  
LIPID 2  
ACT\_SITE 1695  
ACT\_SITE 1709  
VARIANT 223  
VARIANT 228  
VARIANT 376  
VARIANT 643  
VARIANT 669  
VARIANT 726  
VARIANT 726  
VARIANT 2064  
SEQUENCE 2193 AA: 24483 MW: 1143429 Da pI 5.07  
Query Match 5.2% Score 65.5 Fr 1 Length 2193  
Best Local Similarity 19.7% Pred No. 12  
Matches 80; Conservative 50; Mismatches 131; Indels 145; Gaps 18;  
29 TNAVETI FSPFH GAI NADH GAA GNPFAV-----ISIRPMER-- 72  
1828 LEDAV-----YDFHIFALILITSAKYVAI LKKPLLSKIKLILKKIKMVKV 1860  
73 -----KPEH NTELELYVANT PGNPTLPK--IISFV 197  
1881 LNFNVLYVALLKSKKVKAKRSPLIASSI NISVAMPGLFNTYFPHLNDIVTNSA 1940  
108 IGKCEEPVGVKPGSPGFGRLFPKNSLHVSYSVHMLTAP-----KGL 155  
1941 VG--CD-----PDLFSKIPVMTGLHIAHNSGVASLSVWPAVLKLEKLY 1989  
156 TSGKGLAIKNGKTYISKTSPPVVEALSLQFEDP-----TMLNALS 198  
1960 THPE-----TNYID-----YLVNSHHYDPKHFVGRAMPSCGSGTIFNSM 2032  
199 QEVENQ--MLLILG-----PQSGPSLMQSGFELLAMALIEL--VSGLL--DEDEL 249  
2033 NNITPIMKVVYGGHICDGRHMLAVGQNVYASYLWLTASLTALACKRYGLIMTAIRG 2092

250 LFN-----SYFATF-----VHIVL-----PSTLTHHDPND 285  
2693 ECFNEVNTVNTFLKPEFALECYFELVHVMIMKILHESIMVKIKRKHGWSLCL 2192  
286 SVF--MGE--NHWGGEF--KVFAT LGLISN--GFLMLKLYNK 329  
2153 AWNSGPEYFEFFTKPSVPCVGLTLPAST-----LPPKMLDSF 2193

RESULT 4  
A11 METUA STANDARD; PRI: 375 AA.  
ID AAT1 METUA  
AC 060317;  
DT 01-NOV-1997 (rel. 35, Last sequence update)  
DT 01-NOV-1997 (rel. 35, Last annotation update)  
DT 30 MAY-2000 (rel. 39, Last annotation update)  
SF PPGHARF ASPARTATE AMINOTRANSFERASE 1 (EC 2.6.1.1) (TRANSAMINASE A)  
DN (ASPA).  
GN MD0001.  
OS Methanococcus jannaschii.  
OC Archaea: Euryarchaeota: Methanococcales: Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=JAL-1 / DSM 4661 / ATCC 4567;  
PX MENDLINE=9534/9999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald J.M., Clayton P.A., Gocayne J.D.,  
RA Kere J., Rasko J.E., White P., Adams M.P., Belch G.L.,  
RA Grotbeck P., Kirkness E.P., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Smit J.T., Geoghegan N.S.M., Weidman J.F., Fujimaru J.L., Nguyen D.,  
RA International P., Kelley T.M., Peterson J.N., Sadow P.W., Hanna M.C.,  
RA Klotzel M.D., Fellers K.M., Hartz M.A., Kaine R.F., Kodlysky M.,  
RA Klock H.-P., Fraser C.M., Smith H.O., Weiss R.A., Venter J.C.;  
RA Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.  
RT Science 273:1058-1073(1996).  
PL 1- CATALYTIC ACTIVITY: L-ASPARTATE + 2-OXOGLUTARATE -> OXALACETATE +  
L-GLUTAMATE (BY SIMILARITY).  
CC 1- CATALYTIC ACTIVITY: L-ASPARTATE + 2-OXOGLUTARATE - > OXALACETATE +  
L-GLUTAMATE (BY SIMILARITY).  
CC 1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
AMINOTRANSFERASES.  
CC This Swiss-Prot entry is copyrighted. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation.  
CC The European Bioinformatics Institute. There are no restrictions on its  
CC use by non profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.embnet.org/licenses/>  
CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL: U67454; AAR07984.1  
DR TIGR: M09001;  
DR InterPro: IPR001511;  
DR Pfam: PF00165; aminotran\_1.1  
DR Prosite: PS00105; AA\_TRANSFER\_CLASS\_1.1  
KW Transferase; Aminotransferase; Pyridoxal phosphate;  
FT BINDING 223  
FT SEQUENCE 475 AA: 42496 MW: 41498 Da pI 5.64  
Query Match 5.1% Score 94.5 Fr 1 Length 375  
Best Local Similarity 23.7% Pred No. 19  
Matches 47; Conservative 36; Mismatches 60; Indels 55; Gaps 10;

152 PKRTTSRPOIAIKNGKTYISKTSPPVVEALSLQFEDP-----TMLNALS 210  
38 PKHIEAKRALDEGKTHYSPNNGIPELFREISNKKLDDYNL-----DYDKDMLIV-- 88  
211 TSPGSGPSLMQSGFELLAMALIEL--VSGLL--DEDEL 249

DB 89 -----HCAFAFAMLSMTI IDP5--DEVILNPSFVSFTEFAEKKIR- 132  
 QY 271 DAKSTPLHETEDPDSVEMQNNKRWGKEFTVPAFTPIISNGCFEIMKLYK- 328  
 DB 133 -----NIDLEDFENFIDFKVKEFI-----TKKTKLI-IFNSP-SNDTG-----KVKIKET 175  
 QY 429 -----FTHLVSD 336  
 DB 176 LKGLAEIAHDYNNLIVSD 193  
 RESULT 5  
 P010: SVGVH STANDARD: PRT: 2185 AA.  
 AC P13900: Q84794: Q84795: Q84796: Q84797: Q84798: Q84799: Q84800:  
 AN Q84801: Q84802: Q84803: Q84804:  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE GENOME POLYPROTEIN [CONTAINS: COAT PROTEINS VP1 TO VP4; CORE PROTEINS P2A TO P2C; P3A; GENOME-LINKED PROTEIN VP0; PICORNAIN 3C (EC 3.4.22.28) (PROTEASE 3C) (P3C); RNA-DIRECTED RNA POLYMERASE P3D (EC 2.7.7.48)].  
 DE Swine vesicular disease virus (strain UKC/27/72).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
 NC NCBI\_TaxID:12077;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-90364770; PubMed-2168111;  
 RA Steinhilber P., Knowles N.J., McCauley J.W.;  
 RT "The complete nucleotide sequence of a pathogenic swine vesicular disease virus."  
 RL Virus Res. 16:255-274(1990).  
 CC -1- FUNCTION: IT IS THOUGHT THAT THE P2C PROTEIN ATTACHES TO VESICULAR MEMBRANES AND IS ASSOCIATED WITH VIRAL RNA SYNTHESIS.  
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/V SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -1- SUBUNIT: THE VIRUS CAPSID IS COMPOSED OF 60 ICOSAHEDRAL UNITS, EACH OF WHICH IS COMPOSED OF ONE COPY EACH OF PROTEINS VP1, VP2, VP3, AND VP4.  
 CC -1- PPM: SPECIFIC ENZYMATIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3  
 CC This SWISS-PRO entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL collaboration the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X54521; CAA38377.1;  
 DB PIR: A30061; GNNYSY.  
 DB PIR: S11670; S11670.  
 DB HSSP: P03314; ICOW.  
 DB MEKOPS: C03.020;  
 DB INTERPRO: IPR000081;  
 DB INTERPRO: IPR000199;  
 DB INTERPRO: IPR006055;  
 DB INTERPRO: IPR001295;  
 DB INTERPRO: IPR001676;  
 DB INTERPRO: IPR002527;  
 DB Pfam: PF00548; Cys-protease-3C; 1.  
 DB Pfam: PF00947; Pico\_P2A; 1.  
 DB Pfam: PF01552; Pico\_P2B; 1.  
 DB Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DB Pfam: PF00910; RNA\_helicase; 1.  
 DB Pfam: PF00073; rhv; 3.  
 KW Polyprotein; Coat protein; Core protein; Transposase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 FT CHAIN 2 69 COAT PROTEIN VP4 (P1A)

FT CHAIN 70 320 COAT PROTEIN VP2 (P1B).  
 FT CHAIN 331 568 COAT PROTEIN VP3 (P1C).  
 FT CHAIN 569 851 COAT PROTEIN VP1 (P1D).  
 FT CHAIN 852 1001 CORE PROTEIN P2A (P2-3A).  
 FT CHAIN 1002 1109 CORE PROTEIN P2B (P2-3B).  
 FT CHAIN 1101 1429 CORE PROTEIN P2C (P2-3C).  
 FT CHAIN 1430 1518 CORE PROTEIN P3A (P3-1A).  
 FT CHAIN 1519 1540 CORE PROTEIN P3A (P3-1B).  
 FT CHAIN 1541 1723 GENOME-LINKED PROTEIN VP0 (P3B).  
 FT CHAIN 1724 2185 PICORNAIN 3C.  
 FT CHAIN 1541 1723 RNA-DEPENDENT RNA POLYMERASE P3D.  
 FT ACT SITE 1687 1687 MYRISTATE (BY SIMILARITY).  
 FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2185 AA. 23363 MM. C81G38C2334E1DE GRG64;  
 Query Match 5.1%; Score 94.5; PB 1; Length 2185;  
 Best local similarity 19.5%; Freq. No. 26;  
 Matches 78; Conservative 54; Mismatches 136; Indels 131; Gaps 17;  
 QY 29 TENAVETLSPEDRLQALNADCGAAGCTPAAV-----ISTKKMKK- 72  
 DB 1620 TENAV-----VETPPIPAIDLTTSAGYPPYVALGIRKPEIISKRTIDILKIKPMKYG 1872  
 QY 73 -----KREINCOLDELQVYLMDDGNDPDLFKG---LSSEV 107  
 DB 1873 LNI PMATYVNIPLSAIKKAKKSKSPLIFASSTNISSAMGLTQGN YKTHPLNPGIVTSSA 1942  
 QY 108 LGNKEEVPYVMQVGSSEHGLPEPNSLHLYNSVSYVMILTYAKGLISFEGLA 162  
 DB 1933 VG-CD-----PDVEFSKIPVMIDGLIAFDYSGYIASISPMWFTCLKLLKELGY 1981  
 QY 163 LNKGLIYISKISIPVVEAVLISQFHEDF-----IMLNANSQVAVNG 205  
 DB 1982 TNEFTNITP-----YINSSEH YRDPYHVVQZGSCSCSIFNSKTNMILFT 2031  
 QY 206 GMLVLDG-----EGCGEPGMSGCPWELLAMALIEL-VSQGLI--DEKDLITPN--- 253  
 DB 2092 LMLVAVVQINPGPMIAVGVNIVASVPMLEKELALACGAYGCLIMTACGCTETNVT 2091  
 QY 254 IPSVPASLE-----VKQIVP-----KQGSTIDILKICITELSSVMQDN 292  
 DB 2092 WYNTVTLKPEFPADEQYEVLEFVMPMPCTHESLEMTYIPRTGCGVSTCLAMWIRGH 2151  
 QY 293 -ERWAPGERFVVPATFETISMGCFEIMKLYK- 329  
 DB 2152 EYEFIKIPSVVQPTSLPAPST-----LPPWMLDSF 2185  
 RESULT 6  
 ID C81E.GLYEC STANDARD: PRT: 499 AA.  
 AC P93147;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DE CYTOCHROME P450 C81E (EC 1.14.-.-) (ISOPRENOGENE 2'-HYDROXYLASE) (P450 91A4) (CYP 8E-3).  
 DE 91A4 (CYP 8E-3).  
 GN CYP81E1 OR CYP91A4.  
 OS Glycyrrhiza echinata (Licorice).  
 OC Eukaryota; Viridiplantae; Umbellifera; Tracheophyta; Spermatophyta;  
 OC Malvaceae; Glycyrrhizaceae; Glycyrrhizaceae; Glycyrrhizaceae; Glycyrrhizaceae;  
 OC Fabaceae; Papilionoideae; Glycyrrhiza.  
 NC NCBI\_TaxID:46348;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Akashi T., Aoki T., Kameya N., Nakamura I., Ayabe S.-I.;  
 RT "Two new cytochrome P450 cDNAs from alfalfa-induced licorice (Glycyrrhiza echinata L.) cells."  
 RL (in) Plant Gene Register PGR97-167.  
 RN [2]  
 FT CHARACTERIZATION.  
 RX MEDLINE-96099233; PubMed-9706908;





ID E2F5\_MOUSE STANDARD: PRT: 335 AA.  
 AC 061502;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE TRANSCRIPTION FACTOR E2F5 (E2F-5).  
 GN E2F5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Cladara; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclerogathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 PN [1]  
 PP SEQUENCE FROM N.A.  
 PX MEDLINE=95349934; PubMed=7542760;  
 RA Buck V., Allen K.E., Soerensen T., Rybe A., Hijnans E.M.,  
 RA Verhoeven P.M., Petrij F., La Thangue N.H.;  
 RT "Molecular and functional characterization of E2F-5, a new member of  
 the E2F family.";  
 RL Oncogene 11:31-38(1995).  
 RN [2]  
 RP DEVELOPMENTAL EXPRESSION:  
 RX MEDLINE=94025478; PubMed=93976316;  
 RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,  
 RA Phillips F.A.;  
 RT "Expression patterns of the E2F family of transcription factors during  
 mouse development: evidence for a novel E2F family member."  
 RL Mech. Dev. 66:13-25(1997).  
 RN [3]  
 RP DEVELOPMENTAL EXPRESSION:  
 RX MEDLINE=97294922; PubMed=9149906;  
 RA Dagnino L., Fry C.J., Bartley S.M., Farnham P., Gallie B.L.,  
 RA Phillips F.A.;  
 RT "Expression patterns of the E2F family of transcription factors  
 during murine epithelial development.";  
 RL Cell Growth Differ. 8:553-563(1997).  
 RL CELL FUNCTION: TRANSCRIPTIONAL ACTIVATOR THAT BINDS TO E2F SITES. THESE  
 SITES ARE PRESENT IN THE PROMOTER OF MANY GENES WHOSE PRODUCTS ARE  
 INVOLVED IN CELL PROLIFERATION. MAY MEDIATE GROWTH FACTOR-  
 INITIATED SIGNAL TRANSDUCTION.  
 CC -1- SUBUNIT COMPONENT OF THE DRP1/E2F TRANSCRIPTION FACTOR COMPLEX.  
 CC BINDS COOPERATIVELY WITH DP-1 TO E2F SITES. INTERACTS  
 CC PREFERENTIALLY WITH RETINOBLASTOMA-RELATED PROTEINS P130  
 CC AND TO A LESSEER EXTENT WITH P107, THAT INHIBIT THE E2F  
 CC TRANSACTIVATION DOMAIN.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- DEVELOPMENTAL STAGE: IN THE DEVELOPING EMBRYOS. FIRST DETECTED  
 CC IN 13.5-14.5 DPC EMBRYOS. WITH THE APPEARANCE OF STRATIFIED  
 CC EPITHELIUM, LEVELS OF E2F-5 EXPRESSION INCREASE AND BY 16.5 DPC,  
 CC HIGH EXPRESSION FOUND IN THE SUPRABASAL CELL LAYERS. HIGH  
 CC EXPRESSION ALSO FOUND IN OTHER REGIONS WITH STRATIFIED SQUAMOUS  
 CC EPITHELIA INCLUDING THE DEVELOPING PALATE, LIP AND TONGUE. IN THE  
 CC DEVELOPING NERVOUS SYSTEM, FIRST DETECTED IN THE FOREBRAIN AT 9.5  
 CC DPC. AT DAY 10.5, STRONGLY EXPRESSED IN THE POSTAL PRON OF THE  
 CC SPINAL CORD. BY DAY 11.5, E2F-5 IS EXPRESSED THROUGHOUT THE  
 CC DEVELOPING CENTRAL NERVOUS SYSTEM. IN 12.5-15.5 DPC EMBRYOS,  
 CC EXPRESSION FOUND IN THE DIFFERENTIATED VENTRICULAR REGIONS OF  
 CC THE BRAIN. IN THE RETINA, EXPRESSED. IN 14.5-18.5 DPC EMBRYOS, IN  
 CC THE RETINOLASTIC CELL LAYER. IN OTHER DEVELOPING TISSUES, HIGHLY  
 CC EXPRESSED IN THE CHOROIOID PLEXUS. ALSO FOUND IN THE KIDNEY, LIVER,  
 CC LUNG, HEART AND WEAKLY, IN DEVELOPING SPINAL MUSCLE AND  
 CC CHONDROCYTES.  
 CC -1- SIMILARITY: BELONGS TO THE E2F/DP FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.slb.ch/autonomy/>  
 CC or send a email to: [license@isb.slb.ch](mailto:license@isb.slb.ch))  
 CC -----  
 CC EMBL: X86925, CAA60768.1,  
 CC GDB: MG1:105091: E2F5,  
 CC

KW Transcription regulation; Activator, DNA-binding, Nuclear protein.  
 FT DNA BIND 37 108  
 FT DOMAIN 66 88  
 FT DOMAIN 71 108  
 FT DOMAIN 199 405  
 FT DOMAIN 277 335  
 FT DOMAIN 312 329  
 FT DOMAIN 223 226  
 FT SIGNAL 46 84 MW: 46775 Da (46775.00) (P0794)  
 SU  
 Query Match 5.08; Score 92; DR 1; Length 335;  
 Best Local Similarity: 23.5%; Pred. No: 26; Index 86; Gaps 15;  
 Matches 79; Conservative 56; Mismatches 135; Indels 86; Gaps 15;  
 QY 4 MSRSSEVAVN--STLVVAVSMAQPA-----LNAVETIESKPH-----LALINAD 50  
 DB 33 RSPQSSPHEKSLGLTKFVSLDQGGVILKAAATFGLVQPKPRITVITVETP 92  
 QY 51 L-----QAAQRTFAVISTKRM-----MKKRTINQVITELVYINILP 92  
 DB 93 FTRKSNSTQMKVAVNNTKRVITPLVTKAFTFGLTKKELEQAKIWIQSSIXVM 152  
 QY 93 GNENILEKSLSEVLEKNC-----EEVPTVMVPSSEKPLEPPNSL 136  
 DB 153 EDSINNFVSVHIEGLNCHRDITLALVALSVLENTHEMDR-----KRYGI 205  
 QY 137 ELVHSSYSVHLIQAKGLISERGLALNKQITYLSISPVVAVLSQFHEDTTFINA 196  
 DB 206 NIKHSVTHIVL-----SSKRVVFPVTP-----DITQSSQ 245  
 QY 197 KSGPVVAVNMTVITFPGQSTPEKMSCTFWELANAVELVSGRIIEKDKITNI 254  
 DB 246 SSVSVTPKSTMAQNLVQVHVSERS-----QFVQVPAEVSQSSISDILIDELMS 297  
 QY 255 DEYFSELEAVKLVVERDPSFTIDHESF-DLSVEN 289  
 DB 298 SVEFII-RISPIAFQYNNILFNGVGVDFVQI 332  
 RESULT 10  
 PBL2\_PHOSH  
 ID PBL2\_PHOSH STANDARD: PRT: 459 AA.  
 AC P29278;  
 DT 01-DEC-1992 (Rel. 24, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE PHOSPHO-RISPHOSPHATE CARBOXYLASE LARGE CHAIN 2 (PC 4.1.1.39)  
 GN (PHOSFO LARGE SUBUNIT).  
 OS Rhodospirillum rubrum (Rhodospirillum rubrum).  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillum rubrum.  
 OX NCBI\_TaxID=1063;  
 PN [1]  
 PP SEQUENCE FROM N.A.  
 PX MEDLINE=9404861; PubMed=1919998;  
 RA Wagner S.J., Stevens S.E., Nixon B.T., Lambert D.H.,  
 RA Quilley P.G., Jr., Tabita F.R.;  
 RT "Nucleotide and deduced amino acid sequence of the phosphatase  
 PT spheraoides gene encoding form II ribulose-1,5-bisphosphate  
 PT carboxylase/oxygenase and comparison with other deduced forms I and II  
 RT sequences.";  
 RL FEBS Microbiol. Lett. 55:217-222(1988).  
 RN [2]  
 RP SEQUENCE OF 1-36 FROM N.A.  
 RX MEDLINE=9404861; PubMed=1919998;  
 RA Chen T.H., Gibson T.L., McQuie I.A., Tabita F.R.;  
 RT "Identification, expression, and deduced primary structure of  
 PT ribulose-1,5-bisphosphate carboxylase and other enzymes encoded within the form II 202  
 PT fixation operon of Rhodospirillum rubrum.";  
 RL J. Biol. Chem. 269:26447-26454(1994)  
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF  
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC





```

DR Pfam: PF00947: p1co_p2a: 1
DR Pfam: PF01552: p1co_p2b: 1
DR Pfam: PF00680: pna_dep_pna_pol: 1
DR Pfam: PF00910: fna_hel_hase: 1
DR Pfam: PF00073: rnv: 3.
KM RNA-directed RNA polymerase, RNA polymerase, Myristate,
RNA-directed RNA polymerase, hydrolyase, thiol protease, no structure.
FT CHAIN 2
FT CHAIN 70
FT CHAIN 333
FT CHAIN 570
FT CHAIN 571
FT CHAIN 851
FT CHAIN 1001
FT CHAIN 1002
FT CHAIN 1100
FT CHAIN 1101
FT CHAIN 1429
FT CHAIN 1430
FT CHAIN 1519
FT CHAIN 1541
FT CHAIN 1723
FT CHAIN 1724
FT ACT_SITE 1687
FT ACT_SITE 1701
FT CONFLICT 16
FT CONFLICT 177
FT CONFLICT 177
FT CONFLICT 469
FT CONFLICT 469
FT CONFLICT 487
FT CONFLICT 510
FT CONFLICT 516
FT CONFLICT 516
FT CONFLICT 546
FT CONFLICT 546
FT CONFLICT 549
FT CONFLICT 550
FT CONFLICT 554
FT CONFLICT 854
FT CONFLICT 854
FT CONFLICT 873
FT CONFLICT 1097
FT CONFLICT 1097
FT CONFLICT 1290
FT CONFLICT 1437
FT CONFLICT 1437
FT CONFLICT 1503
FT CONFLICT 1503
FT CONFLICT 1616
FT CONFLICT 1616
FT CONFLICT 1624
FT CONFLICT 1624
FT CONFLICT 1627
FT CONFLICT 1630
FT CONFLICT 1630
FT CONFLICT 1718
FT CONFLICT 1718
FT CONFLICT 1734
FT CONFLICT 1734
FT CONFLICT 1758
FT CONFLICT 1758
FT CONFLICT 1824
FT CONFLICT 1824
FT CONFLICT 1867
FT CONFLICT 1867
FT CONFLICT 1880
FT CONFLICT 2001
FT CONFLICT 2001
FT CONFLICT 2095
FT CONFLICT 2095
FT CONFLICT 2115
FT CONFLICT 2115
FT CONFLICT 2175
FT CONFLICT 2175
FT CONFLICT 2179
FT CONFLICT 2179
SQ SEQUENCE 2185 AA: 24451 MW: 1853704.739876 CEC64.

Query Match: 5.0% Score 91.5, DB 1, Length 2185,
Best Local Similarity 19.8% Pred No 355,
Matches 79, Conservation 51, Mismatches 125, Gaps 145,

```

```

DR Pfam: PF00947: p1co_p2a: 1
DR Pfam: PF01552: p1co_p2b: 1
DR Pfam: PF00680: pna_dep_pna_pol: 1
DR Pfam: PF00910: fna_hel_hase: 1
DR Pfam: PF00073: rnv: 3.
KM RNA-directed RNA polymerase, RNA polymerase, Myristate,
RNA-directed RNA polymerase, hydrolyase, thiol protease, no structure.
FT CHAIN 2
FT CHAIN 70
FT CHAIN 333
FT CHAIN 570
FT CHAIN 571
FT CHAIN 851
FT CHAIN 1001
FT CHAIN 1002
FT CHAIN 1100
FT CHAIN 1101
FT CHAIN 1429
FT CHAIN 1430
FT CHAIN 1519
FT CHAIN 1541
FT CHAIN 1723
FT CHAIN 1724
FT ACT_SITE 1687
FT ACT_SITE 1701
FT CONFLICT 16
FT CONFLICT 177
FT CONFLICT 177
FT CONFLICT 469
FT CONFLICT 469
FT CONFLICT 487
FT CONFLICT 510
FT CONFLICT 516
FT CONFLICT 516
FT CONFLICT 546
FT CONFLICT 546
FT CONFLICT 549
FT CONFLICT 550
FT CONFLICT 554
FT CONFLICT 854
FT CONFLICT 854
FT CONFLICT 873
FT CONFLICT 1097
FT CONFLICT 1097
FT CONFLICT 1290
FT CONFLICT 1437
FT CONFLICT 1437
FT CONFLICT 1503
FT CONFLICT 1503
FT CONFLICT 1616
FT CONFLICT 1616
FT CONFLICT 1624
FT CONFLICT 1624
FT CONFLICT 1627
FT CONFLICT 1630
FT CONFLICT 1630
FT CONFLICT 1718
FT CONFLICT 1718
FT CONFLICT 1734
FT CONFLICT 1734
FT CONFLICT 1758
FT CONFLICT 1758
FT CONFLICT 1824
FT CONFLICT 1824
FT CONFLICT 1867
FT CONFLICT 1867
FT CONFLICT 1880
FT CONFLICT 2001
FT CONFLICT 2001
FT CONFLICT 2095
FT CONFLICT 2095
FT CONFLICT 2115
FT CONFLICT 2115
FT CONFLICT 2175
FT CONFLICT 2175
FT CONFLICT 2179
FT CONFLICT 2179
SQ SEQUENCE 2185 AA: 24451 MW: 1853704.739876 CEC64.

Query Match: 5.0% Score 91.5, DB 1, Length 2185,
Best Local Similarity 19.8% Pred No 355,
Matches 79, Conservation 51, Mismatches 125, Gaps 145,

```

Query Match 5.0%; Score 91.5; DB 1; Length 2185;  
 Best Local Similarity 19.5%; Pred. No. 35;  
 Matches 78; Conservative 53; Mismatches 137; Indels 141; Gaps 17;

29 LENVETLSRSHLQALNADGCAAGNTFAV ..... ISTRKMER.. 72  
 1820 LERAV .....  
 73 ..... KRELNQCLEQYVINDLPGDNFLFKG---LSSEV 107  
 1873 LNPMTVYVKELELSADKVKAKSPLEASSINDSAMQTEGCLYKPHI NQGVTSQA 1922  
 108 LKCKCEVPPVWGVGSPGSHGLEPPNSLHVVSSVVMILQAKLTRECA----- 162  
 1933 VG---CD-----PDVWWSKIPVMDLHLAFDQSDASLSFWETLKLTLEKGY 1981  
 163 LNKRIKVIKSTPPVREAVLSQHEDE-----TWELNARSQEVNNG 205  
 1982 TNEFTNID-----YLNSHILYEDKRIYVWQJMSQSSSTFNMINLILFT 2037  
 206 GAVILLIG-----KQSNPSKMSQCFWELLAMAIATLVSQCLL--DEKIDTFN--- 253  
 2032 LMKVVKQIDLCQFMIAVGEDVLAQVEMPIASLAEAGSGYGLMIFAKGGEFNEVT 2091  
 254 -----LPSYFASIFP-----WDLVE-----PMSFPIHIEPQSDSEVENGE 292  
 2092 WIVNTLEKRTFRADQPTLVHVPVMKDIHSIKTKPKPTQHVSLCLLAWHNGEH 2151  
 293 --DKWVGEKRTVRAATFPIISMGPELMKLYDKF 329  
 2152 EYETFKIKIPIVAVCEGISLAFST ...LREKWLDSF 2185

RESULT 13  
 ID PUR2.CHITE STANDARD; PRT: 1371 AA.  
 AC Q26255;  
 DT 01-NOV-1997 (rel. 35, created)  
 DT 01-NOV-1997 (rel. 35, last sequence update)  
 DI 30-MAY-2000 (rel. 39, last annotation update)  
 DE TRIFUNCTIONAL PURINE BIOSYNTHETIC PROTEIN ADENOSINE-3 [INCLUDES:  
 PHOSPHORIBOSYLAMINE--GLYCINE LIGASE (EC 6.3.4.13) (GARS) (GLYCINAMIDE  
 KINOSUCCEPTIDE SYNTHETASE) (PHOSPHORIBOSYLGLYCINAMIDE SYNTHETASE);  
 PHOSPHORIBOSYLGLYCINAMINE CYCLIC-LIGASE (EC 6.3.3.1) (AIRS)  
 (PHOSPHORIBOSYL-AMINOIMIDAZOLE SYNTHETASE) (AIR SYNTHASE);  
 PHOSPHORIBOSYLGLYCINAMINE FORMYLTANSFERASE (EC 2.1.2.2) (GART) (GAR  
 TRANSFORMYLASE) (5'-PHOSPHORIBOSYLGLYCINAMIDE TRANSFORMYLASE)]  
 GN GART.  
 OS Chironomus tentans (Midge).  
 OT Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OT Pterygota; Neoptera; Endopterygota; Diptera; Nemotocera;  
 OT Chironomidae; Chironomidae; Chironominae; Chironomus.  
 OX NRI\_TaxID:7153;  
 RN (1)  
 RP SOURCE FROM N.A.  
 RX MEDLINE=92489368; PubMed=1518084;

HA Clark D.V., Hnikoff S.;  
 FT "Unusual organizational features of the Drosophila Gart locus are not  
 RT conserved within Diptera.";  
 FI J. Mol. Evol. 35:51-59(1992).  
 CC 1 CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYLAMINE + GLYCINE -> ADP +  
 CC ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYLGLYCINAMIDE.  
 CC 1 CATALYTIC ACTIVITY: 10-FORMYLTETRAHYDROFOLATE + 5'-PHOSPHO-  
 CC RIBOSYLGLYCINAMINE + TETRAHYDROFOLATE + 5'-PHOSPHORIBOSYL-  
 CC N-FORMYLTETRAHYDROFOLATE.  
 CC 1 CATALYTIC ACTIVITY: ATP + 5'-PHOSPHORIBOSYL FORMYLTETRAHYDROFOLATE  
 CC ADP + ORTHOPHOSPHATE + 5'-PHOSPHORIBOSYL 5'-AMINOIMIDAZOLE.  
 CC 1 PATHWAY: SECOND, THIRD AND FIFTH STEPS IN DE NOVO PURINE  
 CC BIOSYNTHESIS.  
 CC 1 SIMILARITY: IN THE N-TERMINAL SECTION, BELONGS TO THE GARS FAMILY.  
 CC 1 SIMILARITY: TO OTHER AIRS AND GART FROM BAETIDIA AND EMBRYOTES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/submit/>).  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk).  
 CC EMBL: S43653; AAB23115.1;  
 CC HSSP: P08179; IGAR.  
 DR InterPro: IPR000115;  
 DR InterPro: IPR000728;  
 DR InterPro: IPR001555;  
 DR InterPro: IPR023767;  
 DR Pfam: PF00586; AIRS; 2;  
 DR Pfam: PF01071; GARS; 1;  
 DR Pfam: PF00551; formyl\_transf; 1;  
 DR ProSite: PS00184; GARS; 1;  
 DR ProSite: PS00733; GART; 1;  
 DR MultiFunctional: enzyme, fat line biosynthesis, ligase, transferase.  
 FT DOMAIN 1 433 GARS.  
 FT DOMAIN 434 1171 AIRS.  
 FT DOMAIN 1172 1371 GART.  
 FT ACT\_SITE 1315 1315 BY SIMILARITY.  
 SU SOURCE. 1371 AA; 149103 MW; 7046640d153b9446 CR 64;

Query Match 4.9%; Score 91; DB 1; Length 1371;  
 Best Local Similarity 18.9%; Pred. No. 21;  
 Matches 66; Conservative 56; Mismatches 142; Indels 96; Gaps 17;

41 VASMAQPAI FNAVFIISSPHHLALNADGSAATVIMIAVNSIKKMKKREING- 79  
 529 VAMGVNIVLGNCAAPIGELIV-----IACGHELVITVAIVGADGRKANCA 577  
 80 ---QTLEQVLEN---DLFG---NDFNIFKLSSEVYGNKCEVPPVWGVGSS- 125  
 578 LGGFTAHMISMVGRKYLAVYGVGIIHVELLPLNDVHVG-----VIGLPLSSG 630  
 126 FH-----GRLEPPNSLHV-----HSSYSVHMLTQVAKLTSEKALINKKIYI 170  
 631 IHSNGFSLVNKKIPQGTGFKLTDIAEPSHSIKSYGMEFTPT-----RLVY 675  
 171 SKSPVVEAVLSQHEDETHLNAQSDVPPVNGVWLLKPKQSSPSIMQSTFWEL 230  
 676 STPT-PFTNNGYKALAMITGGGLFRTPTILTRHISVCI ..... DALTKK 720  
 231 LAMAVELVSQGLID EDLETTN ..... HSYVASLPEVKA IYDEKSTPT 275  
 721 LKRVFSMLAAGNVANNEKLPTFNGCIGMIIIMPNDIEMPTDEAMIGISITQDRINCP 780  
 276 LALLEELCLSSVPMQSNKRVNQLKIPVAVLQVLTLSNCTHLKMKK 325  
 781 QVIVKNP--KEVLRKREVIHMKKGDAELISI-SYKQSDVITAGNELVNI 827

RESULT 14



VP3, AND VP4.  
 CC -1- PPM: SPECIFIC ENZYMOLOGIC CLEAVAGES IN VIVO YIELD MATURE PROTEINS.  
 CC CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/PPA IS  
 CC CATALYZED BY P2A. ALL OTHER CLEAVAGES ARE CATALYZED BY P3C  
 CC -1- SIMILARITY: P2A PROTEASE BELONGS TO PEPTIDASE FAMILY 03.  
 CC -1- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY 03.  
 CC This SWISS-PROT entry is copyright © it is provided through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DB EMBL: X80059; CAA56465.1; -  
 DB EMBL: D10582; AAA0149.1; -  
 DB PIR: A36642; GNNYEC.  
 DB MEROPS: C03.001; -  
 DB InterPro: IPR0000081; -  
 DB InterPro: IPR000199; -  
 DB InterPro: IPR000605; -  
 DB InterPro: IPR001205; -  
 DB InterPro: IPR001643; -  
 DB InterPro: IPR001676; -  
 DB InterPro: IPR002527; -  
 DB Pfam: PF00548; Cys-protease-3C; 1.  
 DB Pfam: PF00947; Pico\_P2A; 1.  
 DB Pfam: PF01552; Pico\_P2B; 1.  
 DB Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DB Pfam: PF00910; RNA\_helicase; 1.  
 DB Pfam: PF00073; hmv; 3.  
 DB PRINTS: PR00918; CALICIVIRUSNS.  
 KW Polypeptide; Coat protein; Core protein; Transferrase; Myristate;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease.  
 FT CHAIN 2 69  
 FT CHAIN 70 331  
 FT CHAIN 332 569  
 FT CHAIN 570 861  
 FT CHAIN 862 1011  
 FT CHAIN 1012 1110  
 FT CHAIN 1111 1438  
 FT CHAIN 1440 1528  
 FT CHAIN 1529 1550  
 FT CHAIN 1551 1733  
 FT CHAIN 1734 2195  
 FT LIPID 2  
 FT ACT\_SITE 1697 1697  
 FT ACT\_SITE 1711 1711  
 FT COMPLECT 823 827  
 SQ SEQUENCE 2195 AA; 245405 MW; 10CF5DF288831AF0 CRC64;  
 Query Match 4.9%; Score 90.5; DR 1; Length 2195;  
 Best Local Similarity 19.8%; Pred. No. 43; Mismatches 135; Indels 133; Gaps 18;  
 Matches 79; Conservative 53;  
 QY 29 LENAVETIYPSRPHQAI NAAUIGCAACNTFAV-----ISTIKPMMEK-- 72  
 DB 1830 LKDAV-----YTFETALMLITATYPPYVALGIKKHLLSKPKHLLKIKLKKEMKYG 1832  
 QY 73 -----KREINQUTLELVYVNLDFGDNFTLEKG--LSSEV 107  
 DB 1883 LNLPMVYVVDLPSALUKVAKGKSPILVASSI NISVAMPPI PSLNLYPIHLNPCIIVTCSA 1942  
 QY 108 LCKKCEVPPCYVMGVGSGSEHGLFPRNSILHLVSSSYVMHLIQAP-----KGLISREG 161  
 DB 1943 VG--CG-----PDLFWSKIPVMLDGLIAFDYSGDASLSPVWACILKILLEKGY 1991  
 QY 162 ALNKGKLYISKTSPPVPPAVI SQPHFDP-----TMLNAPSGEVVFN 204  
 DB 1992 L HKEINYID-----YLGNSHHLYPDKHYPFGGMPGSGYSQTSNFMENINNIITR 2040

QY 205 GCVYLILPG-----FGCSPPSGMSQSPFWELIAMAIAEL-VSQGLI--DEKRLDTFN-- 253  
 DB 2041 LMKVYKQIDMGPFPMIAVGNVIVASYPPIIASIACTCKYGLIMFPAIKGQYNEV 2100  
 QY 254 - IFSYFALDE - VKQIYE--FGCSPTDHLHCTVDEVMQ 291  
 DB 2101 LMLNVLLEKRIYALELYYGLVHVPKEMKMLIDESTWELIKRNLGSHVVSCTLAMHNC 2160  
 QY 292 N--LKNVYGEKFTKVVPAIPEIILSNQFCELMKRYKE 329  
 DB 2161 HEVEEFITKTRSPVVGPCI TLPAPST-----LPPWINDSF 2195

Search completed: August 31, 2001, 15:07:34  
 Job time: 372 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Computer Ltd

OM protein - protein search, using sw model

Run on: August 31, 2001, 14:25:47 : Search time 38.18 seconds  
(without alignments)  
710.271 Million cell updates/sec

Title: US-09-577-657A-1  
Perfect score: 1847  
Sequence: 1 FMNNGEESSTVAQSSSTQQ LEAKLPKTSITLIVISKIIS 356

Scoring table:  
BLOSUM62  
Gapop 10 0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1f68:\*  
2: p1f2:\*  
3: p1f3:\*  
4: p1f4:\*

Pred No is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	34.1	371	2 E85430	hypothetical prote
2	570	30.9	359	2 E84513	hypothetical prote
3	504.5	27.3	363	2 E86703	hypothetical prote
4	499.5	27.0	351	2 A86285	protein F91.6 [im
5	496.5	26.9	348	2 T51783	ATP-like protein
6	485.5	26.3	348	2 T51781	protein kinase ACP
7	480	26.0	379	2 T51784	ATP-like protein
8	480	26.0	319	2 T05052	hypothetical prote
9	475.5	25.7	323	2 B86329	hypothetical prote
10	468	25.7	380	2 G64309	hypothetical prote
11	468	25.7	380	2 G64309	hypothetical prote
12	468	25.7	380	2 G64309	hypothetical prote
13	468	25.7	380	2 G64309	hypothetical prote
14	468	25.7	380	2 G64309	hypothetical prote
15	468	25.7	380	2 G64309	hypothetical prote
16	468	25.7	380	2 G64309	hypothetical prote
17	468	25.7	380	2 G64309	hypothetical prote
18	468	25.7	380	2 G64309	hypothetical prote
19	468	25.7	380	2 G64309	hypothetical prote
20	468	25.7	380	2 G64309	hypothetical prote
21	468	25.7	380	2 G64309	hypothetical prote
22	468	25.7	380	2 G64309	hypothetical prote
23	468	25.7	380	2 G64309	hypothetical prote
24	468	25.7	380	2 G64309	hypothetical prote
25	468	25.7	380	2 G64309	hypothetical prote
26	468	25.7	380	2 G64309	hypothetical prote
27	468	25.7	380	2 G64309	hypothetical prote
28	468	25.7	380	2 G64309	hypothetical prote
29	468	25.7	380	2 G64309	hypothetical prote

30	91.5	5.0	2185	1 GNNVSH	genome polypotein
31	91	4.9	550	2 F01770	hypothetical prote
32	91	4.9	2182	1 GNNVBI	genome polypotein
33	90.5	4.9	1374	1 GNNVFC	genome polypotein
34	90	4.9	588	2 T27443	hypothetical prote
35	90	4.9	644	2 T29011	hypothetical prote
36	90	4.9	1123	2 A72311	hypothetical prote
37	90	4.9	2231	2 D71870	conserved hypothet
38	90	4.9	2304	2 T07920	hypothetical prote
39	89.5	4.8	568	2 F71614	probable acetyl-Co
40	89.5	4.8	1032	2 S53571	chromatinic RING f
41	89.5	4.8	1192	2 S55100	hypothetical prote
42	89.5	4.8	2193	2 S52919	hypothetical prote
43	89	4.8	562	2 S27800	polypeptide (1A, 1
44	89	4.8	570	2 T32061	elastase precursor
45	89	4.8	676	2 S41022	gut-specific carbo
					hypothetical prote

## ALIGNMENTS

RESULT 1  
E85430  
hypothetical protein A436470 [unc-f-nd] Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Feb-2001 #sequence: revision 16-Feb-2001 #text-change 16-Feb-2001  
C:Accession: E85430  
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp  
Nature 402, 769-777, 1999  
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.  
A:Cross references: GR-NC\_001068, NID\_97273595, FIDR\_A880113, G3508 G300140  
A:Accession: E85430  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1371 <STOP>  
A:Cross references: GR-NC\_001068, NID\_97273595, FIDR\_A880113, G3508 G300140  
A:Gene: AT9G36470  
A:Map position: 4

Query Match	34.1%	Score 630	DB 2	Length 371
Best Local Similarity	37.3%	Pred No 5.4e-43		
Matches 135	Conservative 84	Mismatches 131	Indels 12	Gaps 5
QY	1	FMNNGEESSTVAQSSSTQQ	LEAKLPKTSITLIVISKIIS	356
DB	11	YMTGDEKTSVAQSSSTQQ	LEAKLPKTSITLIVISKIIS	69
QY	61	AVIST IKMKEKCPRLNQT	ELQVYLNTRFNQNTFKGISEVYQK	111
DB	76	SLIIRIKIVAHNRHIL	IQLEPESFLNDIPGNRFRFKSLPDRHIEIKRINQND	129
QY	112	CEVNVVYVVPSTFTEFLNENSLR	VHSSYSVIMLQVAKELTSEGLAKKQIVIS	171
DB	130	CPSTV FLAAYPSFGRLPENT	THFYVASHSLMLSVPAIDRQKSLINKQVSLIC	187
QY	172	KTSIVVFVAVYSGFRHST	TMILNASEVVRDQCVL	231
DB	188	STSESVASKVNSQFNRHST	FLPCHSEMVSAIMVIL	247
QY	232	AMALIVVSGHIDERTIFNT	TFVASELFEVKI	291
DB	248	SRGAVAVAGTTEFRKLISY	IMFYASATFEDVAKGSPFELPIMFVKKDKNT	307
QY	272	MEAWVQENTFVAVML	ILVLSNDFELMQLYDK	351
DB	306	FDQGVYKAKVAVFVAV	FMVVRGSEKILFLDQVGMVETELARDDIPRTFVVL	357
QY	352	SK 353		
DB	368	FK 369		

```

RESULT 2
F84513
hypothetical protein AL291406 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84513
C:Gene: T23K23.11
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shiu, P.; Bentley, M.L.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Melita, K.S.; Cronin, A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
Cuss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
ATTN: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MIMD-20083487
A:Accession: F84513
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-359 <STD>
A:Cross-references: GB:AE002073; NID:q4388826; PION:AAF19781.1; GSDB:GNC0139
C:Genetics:
A:Gene: T23K23.11
A:Map position: 2

Query Match 30.98; Score 570; DB 2; Length 359;
Best Local Similarity 36.98; Freq. No. 3,4e+38,
Matches 136; Conservative 64; Mismatches 137; Indels 32; Gaps 10;

QY 2 MNRGESSYVQNSSTFOVAVSMAQF-ALDNAVETLSRDFHUALNADLGCAGPNI 60
DB 1 MKGGTGHVATFNSHYORSVYEIQPLVLENVREMLKNGF-CGKVAULGCGTQNTV 59
QY 61 AVISTKMMKKCPENLQTELEQVYLNQ PNDENTLTK-----GLSEVGNKCE 113
DB 60 LAMSAVATVIMESYQSKNPEIDCYLNDPNDENTLTKLFHSFOEKLKPEVKK--- 116
QY 114 FVPCVWGVGSPGRLFPKNSLHVHSSYSVHMTQAPKCLISRECLANKKITYSKT 173
DB 117 ---WFVSGVPSFYSKLPKRSILFVHSAPSIHMLSKIPGLES-----NKSIIKYP 167
QY 174 SPVVKRVAVLSQFHEDETFMELNARSQEVNPGCMVLIIPQAGSDPSDMQCF-TWELLA 232
DB 168 YPSNVKSYLNQFKIPESILFKMPELEVHNGHMLTFVCPKXEDLS-KDCQVMSLIS 226
QY 233 MAIAEISQGLDEKLDITFNISYFASLEVDYKRGSTFTTHIEFT-----LCS 286
DB 227 DCLDLDESGVNDQSKNSMPPYNNEEVEFLIKESFEITIKEDHVPFKYIDR 286
QY 287 VEMQEND-KWVRGEKFTKVRATEPPIISNOGPEIMOKLYKFTHIVSDLE-AKLPK 343
DB 287 EEEDHDSIQLEAFIKHASMAGCTTFPIIVAHGCTAIFPVFVKYAHYAKYISVSNHR 346
QY 344 TTSIIIVLS 352
DB 347 NMIVIVVS 355

RESULT 3
F96703
hypothetical protein T23K23.11 [imported] Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: F96703
C:Gene: T23K23.11
R:Theologis, A.; Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-
Chen, C.W.; Chung, M.K.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Tankins, J.; Johnson-Hopson, C.; Khan, S.; Phaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, P.; Martzall,
Rizzo, M.; Pooney, T.; Powley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MIMD:21016719

```

```

A:Accession: F96703
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-363 <STD>
A:Cross-references: GB:AE005174; NID:q6553892; PION:AAF16558.1; GSDB:GNC0141
C:Genetics:
A:Gene: T23K23.11
A:Map position: 1

Query Match 27.48; Score 594.5; DB 2; Length 363;
Best Local Similarity 33.48; Freq. No. 6,9e+33;
Matches 115; Conservative 73; Mismatches 120; Indels 47; Gaps 8;

QY 2 MNRGESSYVQNSSTFOVAVSMAQFALDNAV-----ETLSRDFHUALNADLGCAG 55
DB 13 MSGTGGPNSYKNSHIOKRTSLKEKIDVLEKNAKLISDSN--TFRIAMGAT 70
QY 56 GPNFAVISTKMMKKCPENLQTELEQVYLNQ PNDENTLTKLFHSFOEKLKPEVK 115
DB 71 GPNTEFLVDNITKSIETLSKNSKPELVFNNDLPNDENTLTSLP-----GDR 122
QY 116 FCIWVSGSFGHGFRTFNSLHVHSSYSVHMTQAFGEISFEDLANKKHYISKTS 175
DB 123 SYLWVGVGSPGRLFPKNSLHVHSSYSVHMTQAPKCLISRECLANKKITYSKT 182
QY 176 PVVPRVATISQFHEDETFMELNARSQEVNPGCMVLIIPQAGSDPSDMQCF-TWELLA 234
DB 183 FVVK-ATPDGPFMRKFTKRAPATEIVSGLLVWGMGILPKMPSNLASIMYISMALV 241
QY 235 IAEVLSQGLDEKLDITFNISYFASLEVDYKRGSTFTTHIEGHIJSEVEMQFNR 294
DB 242 LQHSSELLISEQVDITFNIPYATPEEYTVLVVNGCETV-----EEMELMDTA 293
QY 295 WVRGEKFTK-----VKATEPPIISNOGPEIMOKLYKFT 330
DB 294 WLR--PPTNEDVHVMVCIKATMGSLFINHGELIDVDFRLT 336

RESULT 4
A86285
hypothetical protein T23K23.11 [imported] Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86285
C:Gene: T23K23.11
R:Theologis, A.; Ecker, J.P.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso-
Chen, C.W.; Chung, M.K.; Conway, L.; Conway, A.B.; Conway, A.R.; Greasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Hulzar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Tankins, J.; Johnson-Hopson, C.; Khan, S.; Phaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matli, P.; Martzall,
Rizzo, M.; Pooney, T.; Powley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MIMD:21016719
A:Accession: A86285
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-351 <STD>
A:Cross-references: GB:AE005172; NID:q610411; PION:AAF16558.1; GSDB:GNC0141
C:Genetics:
A:Gene: F9L1.6
A:Map position: 1

Query Match 27.28; Score 499.5; DB 2; Length 351;
Best Local Similarity 33.68; Freq. No. 7,5e+32;
Matches 122; Conservative 73; Mismatches 141; Indels 27; Gaps 10;

QY 2 MNRGESSYVQNSSTFOVAVSMAQFALDNAVETLSRDFHUALNADLGCAGPNI 59
DB 1 MNRGESSYVQNSSTFOVAVSMAQFALDNAVETLSRDFHUALNADLGCAGPNI 60

```







G64300  
 hypothetical protein homolog M10079 - Methanococcus jannaschii  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
 C:Accession: G64300  
 R:Rultz, C.J., Miller, O., Olsen, G.G., Zhou, L., Fleischmann, R.D., Sutton, G.G., Blake, R., Reich, C.T., Ouellet, R., Kikinis, E.E., Weinstock, K.S., Minkley, J.M., Choudhry, A., Sorenson, J.D., Sadow, P.W., Hanna, M.C., Cotton, M.D., Roberts, K.M., Holsclaw, M.A., Science 273, 1058-1073, 1996  
 A:Authors: Kaine, P.; Petrovsky, M.; Kien, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; M010-96337999  
 A:Accession: G64300  
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Features: 1-380 cpm  
 A:Cross references: GI 067463; L77117; R10350001; E100000000; E103500000  
 C:Genetics:  
 A:Map position: FEV73866-72664  
 A:Start codon: TTG

Query Match 57% Score 106, 100% Length 420,  
 Best Local Similarity 20.1% Pred No. 0.78,  
 Matches 64, Conservative 66, Mismatches 127, Indels 62, Gaps 14,  
 Oy 63 ISTIKMKKKPELNGC-----ELQVNLPLSKDELTKLSSEVIGKKEVEVY 118  
 Db 1 MSEIMKLEKIPERINSYFEPFEDIALSTLANE-HTFE----- 41  
 Oy 119 VMGVPGFSEHRI FPPNSIHVHSYVHMTQAPKISPEEL-----ALNKKIYIS 171  
 Db 42 -IGNPGVANSQILRALASH-INANNYERKLTPE-----FTTEDELPLPLSKELKNDPRVR 95  
 Oy 172 KTF--PVVEAVLSQHECTFPLNAR--SGEVINQGVW-----LILKPGKSD 218  
 Db 96 KTSQYPLAEVFLDEVFANSILNALISINERLYHNGDLKRVPLISLPGASNELPE 155  
 Oy 219 PSMQSGPTMELLAMALVELS-VGLDEKLEFNISYFSLSEVNDIVERGSEFTID 277  
 Db 156 ENELAFYDFELFRKVPKIPSCENLVKLIKIDFEYKPTTISIELPKMKDKEANVDIF 215  
 Oy 278 HEFTPL SEVNGEMKWEKPELVVAF-EPISNGGFEIMKTXDKFTIHVVS 335  
 Db 216 NITGVVDIKKKKSNHITISIPREFKSKAIKFAVINGFEAFITF-LETFIHITWD 273  
 Oy 336 DEAKLPKTSILVLSKI 354  
 Db 274 DIDP-----LLIVSKV 284

## RESULT 11

T41720  
 hypothetical protein SPAC821.09 - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T41720  
 R:Rigert, M., Lyut, M., Rajandream, M.A., Barrell, B.G.  
 A:Submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z22612  
 A:Accession: T41720  
 A>Status: preliminary; translated from cDNA/EMBL/GenBank  
 A:Molecule type: DNA  
 A:Residues: 1-1016 cpm  
 A:Cross references: EMBL AL12770; F10N CA57443.1; GSPFB GNG006; SPAC821.09  
 A:Experimental source: strain 972h; csmid 0621  
 C:Genetics:  
 A:Gene: SPAC821.09  
 A:Map position: 1

Query Match 55% Score 101.5, 100% Length 1016,  
 Best Local Similarity 18.4% Pred. No. 5.6,

Matches 70, Conservative 43, Mismatches 127, Indels 139, Gaps 14;  
 Oy 9 SSVAQNSPPEVYASMAQ--PALENAVEPLISP-----DPHQALNMAAD-L 51  
 Db 2 SSVAQNSPPEVYASMAQ--PALENAVEPLISP-----DPHQALNMAAD-L 51  
 Oy 52 GSAAGNIPAVISIKMKKKKRLNVLITGVYINLHGNINILPKG--LSS 105  
 Db 52 GIVSHVILAVSI-----LSSLELKKLEEDNLLSEPLINMAYALTYWMS 112  
 Oy 106 EVLVNMC-----EEVPTVYVCGSPHGLPFPNSIHVHSSVHMTQA 151  
 Db 113 GPMGSGIATHTNGCVMDPTIPSY-----YE 142  
 Oy 152 PKQTSPEGLAKKKTYSKTSPPVVPATISQHEPTMFIANSSGVPPGCKYIT 211  
 Db 143 PHLVTS--LGSASLISNITFIVE--HMSARPT----- 175  
 Oy 212 KCMQSSALSLMSSCEWELLAMALVELS-VGLDEKLEFNISYFSLSEVNDI 271  
 Db 176 -----PSWSSSQMLT--LLEGNAVITAYTNATFOISSILYINDVEVP 219  
 Oy 272 GPTLHHEFTPLSEVMEPEKRW--VPGKPTKVVAFTPEIISNQ----- 317  
 Db 270 QLSAVAKYF--VMSINHWMLLYVGLSTLLETSYLMVSGTFNGYIQAKITL 273  
 Oy 318 GPELMKLUKFTIHVSDI 337  
 Db 274 GPCVLAELYTFVAVYITGI 293

## RESULT 12

T15137  
 hypothetical protein T28F2.3 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 C:Accession: T15137  
 R:Madsen, C.; Promick, B.  
 A:Submitted to the EMBL Data Library, April 1997  
 A:Reference number: T15137  
 A:Accession: T15137  
 A>Status: preliminary; translated from cDNA/EMBL/GenBank  
 A:Molecule type: DNA  
 A:Residues: 1-319 cpm  
 A:Cross references: EMBL AF000109; M103204/141; F103204/141; F10N AAF53554.1; GSPDB  
 A:Experimental source: strain Bristol NC; clone T28F2  
 C:Genetics:  
 A:Gene: CESD:T28F2.3  
 A:Map position: 1  
 A:Introns: 90/2; 189/3; 266/3; 295/3

Query Match 54% Score 99, DB 2, Length 319,  
 Best Local Similarity 20.3% Pred. No. 1.9,  
 Matches 75, Conservative 56, Mismatches 106, Indels 132, Gaps 18,

Oy 1 FNNVSESSAYANSSFTQVYASMAQPA-----LENAVETLSR--DFHL 43  
 Db 3 FLSGLDQ--GFAD-----LKQKALMAFTALSLNQLKRAKPMESAFETGINKKLISLS 56  
 Oy 44 GAIANAALQAAAC--MFAVVISIKMKKKKPELVNVLITGVYINLHGNINILPKG 102  
 Db 57 KVLNVEISLFTGSPIDIVPITAFGEHLNNAHFVETSETFKAVND--GNSITLMBEG 114  
 Oy 103 LSSFVIGNNRHYLV--VMSYVPGSPHGLPFPN-----SLHVVSSVHMT 144  
 Db 115 NSSRLATSPPEQVYHIDAV--NFHMAETPMNGSEHTIGGVYAGEMHLIRN----- 165  
 Oy 150 GAPPALSPPELALNKKIYISKTSPPVVPATISQHEPTMFIANSSGVPPGCKYIT 211  
 Db 166 --TFPAIMAIKQKPSVIAIA-----VFINPSHGNVAVESPLINLPG----- 207

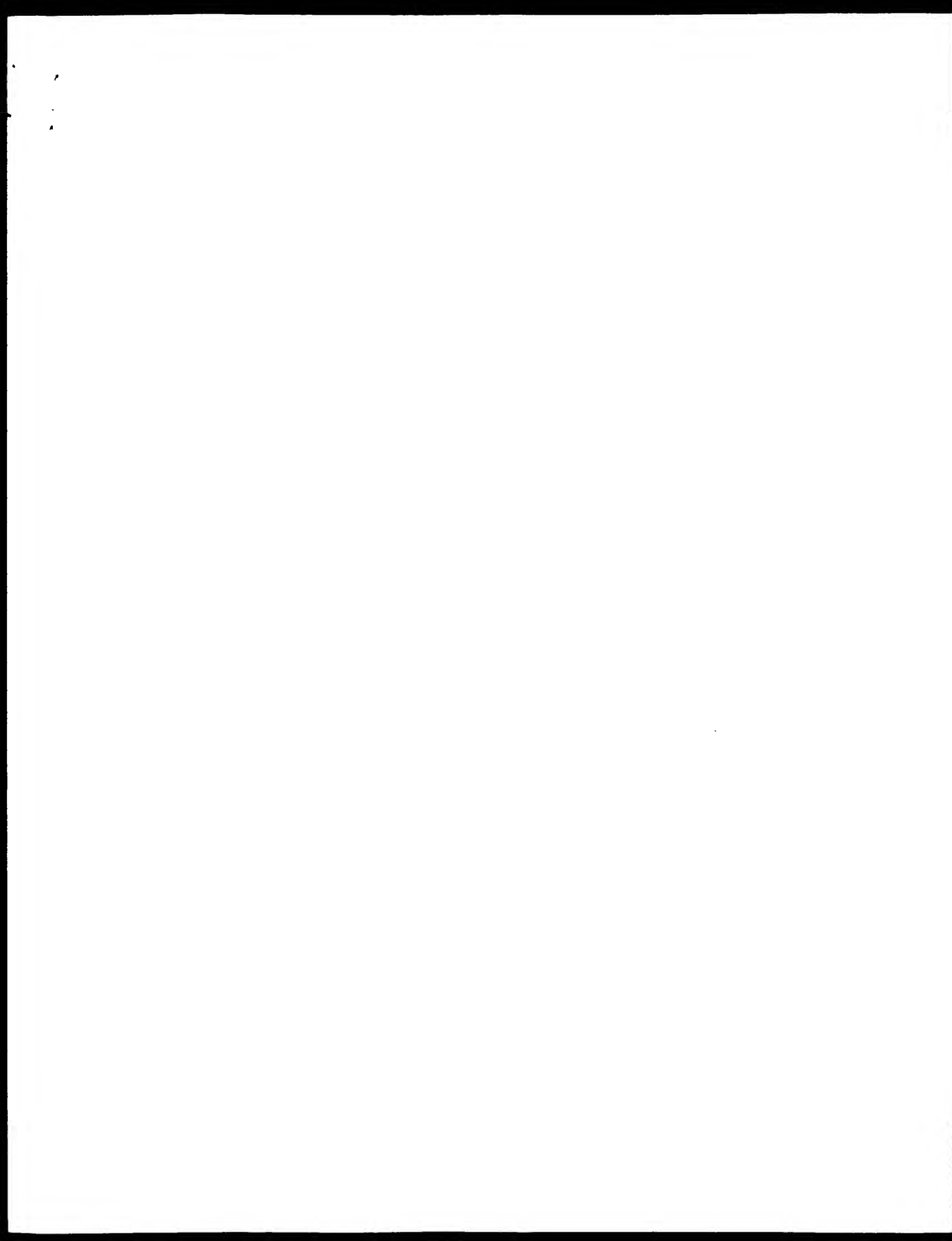


```

OY 29 LNNVETLESFDPHIGALNADLGCAGPNTAV-----ISTKEMMEK-- 72
Db 142A LETAIV-----YVTHGTEFATIGLTSAGYVAIIGIKKRIIISKRIKDIKIKKIMDKYG 1880
OY 73 -----KCREINCOLELQVYLNDLEFQNDENTLFKG---LSSTV 107
Db 1881 LNLPMVTYVDELPSSPKVAKKSPITFASSINDSVAMPQTPGNIYKTFHINPGIVTQSA 1940
OY 108 LKNGCFPPQYVMGVPSSEFHGRLFPENSLHLYSSYSVHMLTQAP-----KGL 155
Db 1941 VG--CD-----PDLFWSKIPVMDGHLAFQVSGYDASISPVWFACTIKLLEKLOY 1989
OY 156 TSPFGIAI NGKRIYISKTSPPVYFPAVISOPEDE-----TMFLNAPS 198
Db 1940 THRP-----TWYID-----YLCNSHLVPPKHYFVGGNPSQSCSTISFNSKI 2032
OY 199 QEVVPGWVLLIHG-----PQSDPSIMQSFTEWELAMAIAL-VSQGLI--DEKRL 249
Db 2033 NNIITPIMKVKYKGTITDQFPMIATCTQVIASTWPTDASIIAEACPVYCIIMTPADKG 2092
OY 250 DTFN-----IPSYFASLIE-----VKDIVE-----PDGSPFTTHIEGFPID 285
Db 2093 FPFNEFTWTVTFIKPYFPADEQYPLVHPVMPMKDIHESIPTKDKPKNTQCHVPSLICL 2152
OY 286 SVE--WGENCKWYCFEFTYVVPATIEPLISNCFSEIMMKLYDKF 329
Db 2153 AMHNGEDYFEPFIPKIPSVVVGCLTLPABST-----LPPKWLDSF 2193

```

Search completed: August 31, 2001, 15:01:17  
 Job time: 2130 Sec



Genetic version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd

CM protein - protein search, using SW model

Run on August 31, 2001, 14:34:27; Search time: 08 seconds  
(without alignments)  
281,065 Million cell updates/sec

Title: US-09-577-657A-1

Perfect score: 1847  
Sequence: 1 FMNRTGSSVAQMSPTQQ LEAKLRKTSIIIVSKIIW 456

Scoring table: BLOSUM62  
Gapop 10 0, Gapext 0 5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: Issued Patents AA:  
1: /cgn2\_6/pdata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/pdata/2/1aa/5H.COMB.pep.\*  
3: /cgn2\_6/pdata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/pdata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/pdata/2/1aa/6C.COMB.pep.\*  
6: /cgn2\_6/pdata/2/1aa/6D.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No	Score	Query Match	Length	DB ID	Description
1	97.5	53	2763	US-08-496-944-2	Sequence 2, Appl
2	85	46	885	US-08-484-105-14	Sequence 14, Appl
3	85	46	885	US-08-484-106-14	Sequence 14, Appl
4	81	44	553	US-09-068-655-6	Sequence 5, Appl
5	80	43	441	US-08-314-309A-4	Sequence 4, Appl
6	79.5	43	118	US-08-545-809A-145	Sequence 145, App
7	79.5	43	2465	US-08-596-291-3	Sequence 3, Appl
8	79.5	43	2465	US-09-100-804-3	Sequence 3, Appl
9	79	43	428	US-07-882-790-4	Sequence 4, Appl
10	79	43	1429	US-08-540-804-14	Sequence 14, Appl
11	79	43	1429	US-08-418-245-14	Sequence 14, Appl
12	79	43	1429	US-08-521-872-14	Sequence 14, Appl
13	79	43	1420	US-08-590-399-14	Sequence 14, Appl
14	78.5	43	2296	US-08-286-819A-27	Sequence 27, Appl
15	78.5	43	2296	US-08-286-819A-27	Sequence 27, Appl
16	78	42	438	US-08-577-492-38	Sequence 38, Appl
17	78	42	968	US-09-180-439-3	Sequence 3, Appl
18	78	42	968	US-09-180-439-3	Sequence 3, Appl
19	78	42	1016	US-09-180-439-8	Sequence 8, Appl
20	78	42	2306	US-08-480-604A-13	Sequence 10, Appl
21	77.5	42	2306	US-08-480-604A-13	Sequence 10, Appl
22	77.5	42	139	US-08-663-566A-9	Sequence 9, Appl
23	77.5	42	501	US-08-023-610-9	Sequence 9, Appl
24	77.5	42	581	US-08-288-062A-9	Sequence 9, Appl
25	77.5	42	581	US-08-288-062A-9	Sequence 9, Appl
26	77.5	42	501	US-08-804-372A-7	Sequence 7, Appl
27	77.5	42	501	US-08-804-372A-7	Sequence 7, Appl

28	77.5	42	501	US-08-10245-9	Sequence 9, Appl
29	77.5	42	1577	US-08-793-834-2	Sequence 2, Appl
30	77	42	435	US-08-111-930-16	Sequence 16, Appl
31	77	42	435	US-08-452-252-2	Sequence 2, Appl
32	77	42	435	US-08-734-550-2	Sequence 2, Appl
33	77	42	435	US-08-459-075-8-2	Sequence 2, Appl
34	77	42	456	US-08-643-036A-4	Sequence 4, Appl
35	77	42	456	US-08-643-036A-4	Sequence 4, Appl
36	77	42	476	US-09-233-989-2	Sequence 2, Appl
37	76.5	41	468	US-07-477-630-4	Sequence 4, Appl
38	76.5	41	720	US-08-840-236-1	Sequence 1, Appl
39	76.5	41	720	US-08-505-448A-1	Sequence 1, Appl
40	75.5	41	124	US-08-428-197-44	Sequence 44, Appl
41	75.5	41	124	US-08-428-197-44	Sequence 44, Appl
42	75.5	41	124	US-08-531-927B-4	Sequence 4, Appl
43	75.5	41	291	US-07-868-357A-12	Sequence 12, Appl
44	75.5	41	354	US-08-497-804-21	Sequence 21, Appl
45	75.5	41	354	US-09-124-807-21	Sequence 21, Appl

## ALIGNMENTS

RESULT 1  
US-08-496-944-2  
Sequence 2, Application US/08496944  
Patent No. 6040496  
GENERAL INFORMATION:  
APPLICANT: Law, Marcus D  
TITLE OF INVENTION: Use of translationally altered RNA to  
TITLE OF INVENTION: Control Resistance to Murine Heart Muscle Virus and Other  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-Geigy Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: NY  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-TS/MS-DOS  
SOFTWARE: PatentIn release #1.0, Version #1.308  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,944  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY, FIRM, INFORMATION:  
NAME: Elmer James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1814  
INFORMATION FOR SEQ ID NO.: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2763 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULAR TYPE: protein  
US-08-496-944-2

Query Match: 534; Score 97.5; EB 3; Length 2763;  
Best Local Similarity 19.88; Pred. No. 1.3; 144; Indels 93; Gaps 17;  
Matches 745; Conservative 58; Mismatches 144;  
5 CDESSVAAGSEFLVYASMAQALERAVE-ILFSEFLHGLATNAAADGCAACHTFAVI 63  
1472 GYKSKMKNAAFLKIDSKVAFIVGSHFLYGLHVAWG ..... FVV 2013  
64 STIRMEKRC...FELNGTLELVYINTEFGMDFTLTKGLSS-----EVINKKEE 114

Db 2014 NLKNVGMQCVVYVDEEFIFRSINI NAAGALYTGKKNVFNENSSDKFEITMESCEP 2073  
 QY 115 VPCVMEV-PGSHGRLFPNLSLH-----VHSYSVH-----146  
 Db 2074 IYKXJGVMGNSJKAETKPEKTMANKTRIFVAPLETLGGVYVCDDENNQYSHLPG 2133  
 QY 147 -MLTGARK-GLTSREGIALNKRY-----ISKTSPEVREAYLS--QFHEPF--190  
 Db 2134 FWVATIKREYGTMRRLLEKLEPGGVYCDAGSQPDSSLTFYLLINAVLNIRKQFMEIWDIG 2193  
 QY 191 -TMEIANSJAY --- VFNQWVI IIPGQCSIPSDMSQCFWELLAMAIAELVSQGLI 244  
 Db 2194 AQMKNYITELVYPIAPIDSTIVKRFKGNNSQGPSTVVDNIIIMVIAFNVA-MLSSG-1 2251  
 QY 245 DEDKLIFFNIHSYASLEEKDIVERGSPFIIDHIE-----GFDLSVEMQENDKM- 295  
 Db 2252 KEHID- NCKMFMANDDLIAVHRIPEFTLDEFGHFNGLNLFETSKIKQKSLMF 2509  
 QY 296 --VKGKPK 301  
 Db 2310 MSTKQIKY 2417

RESULT 2  
 US-08-484-105-14  
 : Sequence 14, Application US/08484105  
 : Patent No. 5589341  
 : GENERAL INFORMATION:  
 : APPLICANT: STILLMAN, Bruce  
 : APPLICANT: BELL, Stephen P  
 : APPLICANT: KOBAYASHI, Ryuji  
 : APPLICANT: RINE, Jasper  
 : APPLICANT: FOSS, Margit  
 : APPLICANT: MCNALLY, Francis J  
 : APPLICANT: LAURENSEN, Patricia  
 : APPLICANT: HERSKOWITZ, Ira  
 : APPLICANT: LI, Joachim J  
 : APPLICANT: GAVIN, Kimberly  
 : TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
 : NUMBER OF SEQUENCES: 24  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
 : STREET: 4 Embarcadero Center, Suite 3400  
 : CITY: San Francisco  
 : STATE: California  
 : COUNTRY: USA  
 : ZIP: 94111-4187  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/484,105  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Osman Ph.D., Richard Aron  
 : REGISTRATION NUMBER: 36,627  
 : REFERENCE/DOCKET NUMBER: A-59032/21F/PAO  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (415) 494-8700  
 : TELEFAX: (415) 494-8771  
 : TELELEX: 910 277299  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:  
 : LENGTH: 885 amino acids  
 : TYPE: amino acid  
 : STRANDEDNESS: single  
 : TOPOLOGY: linear  
 : MOLECULE TYPE: protein  
 : US-08-484-105-14

Query Match 4.68; Score 85; DB 1; Length 885;  
 Best Local Similarity 19.08; Pred. No. 4.4;  
 Matches 60; Conservative 50; Mismatches 95; Indels 110; Gaps 10;

QY 50 DIFGAAGNTFAVIST--IKFMKEKCPETNQTLEIYVINDLPG-----NDENIL 99  
 Db 391 DIDLALFFRRFVYSAKGMETTFESKVKKGLNSFNKSEFVYKAVTQNTLLPARENFFASL 450  
 QY 100 FKGLSSVYIGNKCEVPYMGVPG-----SPHGRLEPR-----NSLH 197  
 Db 451 YLSTYSALTEAGISTSI--YIAGTPGVGKILTVREYVKDMLTSAVKLELPRFGYIEINSLK 508  
 QY 138 LVHS SYSHWMLTGAPKGLTSFGIAIANKRYIKTSIPVYVREAYLSGFHEDETMILN 195  
 Db 509 IVKASUSTEYFMOKISGKGLIS--GAMESLSEYFMKV--PAIKKRPVAVLIDELDA-LV 563  
 QY 196 APSQVVPNPMVWLIPGQCSIPSDMSQCFW--ELAMAIAELVSQGLIDEPKL 249  
 Db 564 SKSQDV-----MYNFMNATYSNAKLIVAVA-----590  
 QY 250 DTFNIPSY--FASL-----HVKQVIV--KQSGTIAATISF--282  
 Db 591 NTLDPERHLGNKLSRIGFTFRIMFTGYTHEELRTIINLRYLNSSSYVYDPTGSSYM 650  
 QY 283 --DLSVEMQENDK 294  
 Db 651 ISPSSTIETDEKPK 665

RESULT 3  
 US-08-484-106-14  
 : Sequence 14, Application US/08484106  
 : Patent No. 5614618  
 : GENERAL INFORMATION:  
 : APPLICANT: STILLMAN, Bruce  
 : APPLICANT: BELL, Stephen P  
 : APPLICANT: KOBAYASHI, Ryuji  
 : APPLICANT: RINE, Jasper  
 : APPLICANT: FOSS, Margit  
 : APPLICANT: MCNALLY, Francis J  
 : APPLICANT: LAURENSEN, Patricia  
 : APPLICANT: HERSKOWITZ, Ira  
 : APPLICANT: LI, Joachim J  
 : APPLICANT: GAVIN, Kimberly  
 : TITLE OF INVENTION: ORIGIN OF REPLICATION COMPLEX GENES  
 : NUMBER OF SEQUENCES: 24  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: FLEHR, HOBACH, TEST, ALBRITTON & HERBERT  
 : STREET: 4 Embarcadero Center, Suite 3400  
 : CITY: San Francisco  
 : STATE: California  
 : COUNTRY: USA  
 : ZIP: 94111-4187  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: floppy disk  
 : OPERATING SYSTEM: IBM PC compatible  
 : SOFTWARE: Patent in Release #1.0, Version #1.30  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/484,106  
 : FILING DATE:  
 : CLASSIFICATION: 435  
 : ATTORNEY/AGENT INFORMATION:  
 : NAME: Osman Ph.D., Richard Aron  
 : REGISTRATION NUMBER: 36,627  
 : REFERENCE/DOCKET NUMBER: A-59032/21F/PAO  
 : TELECOMMUNICATION INFORMATION:  
 : TELEPHONE: (415) 494-8700  
 : TELEFAX: (415) 494-8771  
 : TELELEX: 910 277299  
 : INFORMATION FOR SEQ ID NO: 14:  
 : SEQUENCE CHARACTERISTICS:

LENGTH: 885 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-484-106-14

Query Match 4.8%; Score 85; DB 1; Length 885;  
Best Local Similarity 19.0%; Pred. No. 4.4;  
Matches 60; Conservative 50; Mismatches 95; Indels 110; Gaps 16;

QY 50 UUGAGAGNITAVIST--IKKMKKKKRLNQTLELYLNDLGG-----NDFNLL 99  
DB 391 DLDIALEERPRIVSAKGMETTFISKVKKOLNSKKEETIKKADDPNYPAPENEFASI 450  
QY 100 FKGLSEVTKCKFEVPCYVGVPG-----SPHGRLFP-----NSLH 137  
DB 451 YLSLXSAIEAGTSTSI--YIAGTGVGKTLTVREYVADMTSADQKELPRVYIEINQIK 508  
QY 138 LVHS--SYVHWLTQAKKGLTSREGALNKGITYISKTSPPVVRVAYLSQFHELFIMELN 195  
DB 509 IVKASDSYEVEWQKTSGEKLITS--GAAMESLEEFYFNKV--PATKKRPVVLLEFLNA-LV 563  
QY 196 ARSQEVVNGCMVJLNGRQSDPSDMQSCFTW-----ELLAMATAELYSQGLIEDKIL 249  
DB 564 SKSDGV-----MYNEFNMATYSMAKLIVAVAA----- 590  
QY 250 DTFNIPSY-----FASL-----EYKADIVE-----RDQSFTIDIEGCF-- 282  
DB 591 NTLDLPERHILANKISSRIGTFPIWFTGYTHRLRTILNLEPLKTYLNSSPFVDFPTGSSYM 650  
QY 283 ---DIDSVEFMQPNRK 294  
DB 651 ISPDSSTLEDEER 665

## RESULT 4

US-09-068-655-6  
; Sequence 6, Application US/09068655A  
; Patent No. 6196579  
; GENERAL INFORMATION:  
; APPLICANT: JACKSON, Ronald James  
; TITLE OF INVENTION: METHOD OF PRODUCING ALPHA 1,4-STANDARD TRANSFERASE  
; FILE REFERENCE: 026579-186  
; CURRENT APPLICATION NUMBER: US/09/068,655A  
; CURRENT FILING DATE: 1998-08-31  
; EARLIER APPLICATION NUMBER: AU PN 6587  
; EARLIER FILING DATE: 1995-11-15  
; EARLIER APPLICATION NUMBER: PCT/AU96/00725  
; EARLIER FILING DATE: 1996-11-15  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Myxoma virus  
US-09-068-655-6

Query Match 4.4%; Score 81; DB 4; Length 553;  
Best Local Similarity 20.6%; Pred. No. 5.6;  
Matches 65; Conservative 40; Mismatches 97; Indels 114; Gaps 13;

QY 62 VISITKMKKKKRLNQTLELYLNDLGGNDPNTLFKGLSEVY-----GNKCEVP 116  
DB 144 VITTIHIVSVLRQPNFKILIELCLNLTSSDLNLEEDVCAVYLVWLKNNMELDP 203  
QY 117 CYV-----MGVFGSHGKLFKPNSSLHLYSSYSVWLVQAKGSLKSLGLANNG 166  
DB 204 SVLEQVMSLSLMSVKNLLAKTCTFNKRYV-----QSLAKLIDHSPPPTQGTLSIGGP 259  
QY 167 KIVISKISPV-----VREAVI-----SDFHE-----DFMFLNKKSEVVP----- 203

DB 260 KYDDITSSPELVSPVDWVTVSYLPTHRQFSSVAIDFVYVVGGLQDSVSAVS 319  
QY 294 -----NQMVLIR-----GPGCSDFSDMSQCFWELLAMALA 236  
DB 320 YIVKTNKKKEPLKSRHROGLVILPKLIVISGGRN-----SYLKIVYVW----- 368  
QY 237 ELVSQGIJHEIKLPPNIPSY-----FASLEVKD-----IVEKDSFTIDIEGFCO---- 283  
DB 369 -----PYATWFKLSLREAFITVGAAYFNKYYITGJSVEPSR 410  
QY 284 ---LDSVEMQENDKVV 296  
DB 411 LFTDTVFTQNNKVV 426

## RESULT 5

US-08-314-309A-4  
; Sequence 4, Application US/08314309A  
; Patent No. 5677141  
; GENERAL INFORMATION:  
; APPLICANT: TSUGAI, TAKAO  
; APPLICANT: FUKAGAWA, MASAO  
; APPLICANT: IWAMI, MORITA  
; APPLICANT: ARAMORI, ICHIRO  
; APPLICANT: KUDO, HITOSHI  
; TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: ORION, SPIVAK, MCLELLAND, MATER & NEUSTADT,  
; STREET: 1755 S. Jefferson Davis Highway, Suite 400  
; CITY: Arlington  
; STATE: Virginia  
; COUNTRY: U S A  
; ZIP: 22202  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/314,309A  
; FILING DATE: 30-SEP-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/631,906  
; FILING DATE: 21-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OBLON, NO. 5677141man F.  
; PRTISTRATION NUMBER: 24,618  
; REFERENCE/WORKER NUMBER: 18-863-0 CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 413-3000  
; TELEFAX: (703) 413-2220  
; TELEX: 248855 OPAT DR  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 441 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-314-309A-4

Query Match 4.3%; Score 80; DB 1; Length 441;  
Best Local Similarity 22.0%; Pred. No. 4.9;  
Matches 35; Conservative 30; Mismatches 54; Indels 40; Gaps 7;

QY 139 VHSYSVHWLTQAKKGLTSREGALNKGKITYISKTSPPVVRVAYLSQFHEPDTMFLNARS 198  
DB 25 INLPSPGLHINFNANVILPR-LAVIHGALPAAALAP-----TLQK 64





APPLICANT: HELDIN, CARL-HENRIK  
TITLE OF INVENTION: PRIMARY STRUCTURE AND FUNCTIONAL  
TITLE OF INVENTION: EXPRESSION OF NUCLEOTIDE SEQUENCES FOR NOVEL PROTEIN  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.  
STREET: 600 ATLANTIC AVENUE  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02210  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In P-1-95a- #1 0, Version #1 25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/100,804  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/596,291  
FILING DATE: 09-AUG-1996  
APPLICATION NUMBER: US 08/115,573  
FILING DATE: 01-SEP-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/06943  
FILING DATE: 01-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: GATES, EDWARD R.  
REGISTRATION NUMBER: 31,616  
REFERENCE/BOOKLET NUMBER: J44h1/7003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-720-3500  
TELEFAX: 617-720-2441  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2465 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-100-804-3

Query Match 4.38; Score 79.5; DB 3; Length 2465;  
Host Local Similarity 23.08; Pred. No. 87;  
Matches 46; Conservative 35; Mismatches 91; Indels 37; Gaps 8;

QY 2 MNGDESSYAQSSFTQVYASMAQPALENAVETLESRPHLQALMAALGCAAGPNTFA 61  
DB 966 MSSSYHRTSGASLYPRKKNVYVMEPPQYVAFLVSRPSHMSPSDAESLAQVTKINSK 1025  
QY 62 VISTIKMKMKKRPINCOLE--LQVINDLFQNDN-----TFFK-----GI 103  
DB 1026 SVASLINSRPRKHEDSSSTRPQAYVIVYHKRWSLVSPREDEITLVNKKKAKGL 1085  
QY 104 SSEYIGKCE-----VPCYVWGVGSHG-----RFPNNSLHVHSSY--SVHMIT 149  
DB 1086 GFOIIGSEKKEFTGJFISVAVRQSPAFPHQI KPDQGLSVNSVSLQVSHHALELIQ 1145  
QY 150 QAPKGLSTRGLANKGIYISK-SPV 177  
DB 1146 NAEDEV----LVISQPKRKISKVPSTPV 1170

RESULT 9  
US-07-882-790-4  
Sequence 4; Application US/07802750  
Patent No. 5492695  
GENERAL INFORMATION:  
APPLICANT: BRIEF, ROBERT R

APPLICANT: FRANK, GLENN  
TITLE OF INVENTION: VARIANTS FOR SUSCEPTIBLE HOSTS AGAINST  
TITLE OF INVENTION: NONADAPTED PARASITES  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: MORRISON & FORSTER  
STREET: 755 PAGE MILL ROAD  
CITY: PALO ALTO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94304-1018  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In P-1-95a- #1 0, Version #1 25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/842,790  
FILING DATE: 19920514  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: SETHRA, DEBRA A.  
REGISTRATION NUMBER: 33,309  
REFERENCE/BOOKLET NUMBER: 27010-2000/00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 813-5600  
TELEFAX: (415) 494-0792  
TELEX: 706141  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 428 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-882-790-4

Query Match 4.38; Score 79; DB 1; Length 428;  
Host Local Similarity 21.14; Pred. No. 6;  
Matches 75; Conservative 53; Mismatches 115; Indels 112; Gaps 20;

QY 71 FRKTPRINTTLEFY---VYINDPQIN+NLIS\*-----TSSVYFNQTPVAVVWV 121  
DB 54 KKKAAE-----ELGGGCMALSEIIIGNKKMLKIKISADPQIDKKVEDMLKLVVD 107  
QY 122 -----VPSFPRPFPNPST--HIVSSYSVH--WLTAPKGLTSP---EGIA 162  
DB 108 KKKKFLFEVAVVGFIVYAMNFEFFENEHLESYFTYLSWLTACFTEIKKKEEG-- 165  
QY 163 LNKGIYISKTSPPVVRKAVVLSQHHDFITPLNARSQEVVPGCVLL--LRG----- 213  
DB 166 --KSKMTGKK-----ITVYPSITGKKKRAAEFGSPMAALPVLGEHAKTVI 214  
QY 214 RQCSDFSDMOQCTMLMLAMALAIIVSGLIDED--KIDP----- 252  
DB 215 KKKKKI-----SATAPELKNVKEHMF-KIVYVKKKKKILVAVPQKFAVIVHPPPPN 268  
QY 253 --NIFSYF-----ASLENYKQIVERPDS-----TTHHIGSLD--SVEMQ 290  
DB 269 FHNFSFYQTVI SWITPAKPEIKKKRPPQSKMKQLKKKTVVFEISTGPKKRAAEPIQ 328  
QY 291 ENDKW-----VKGKPTKRVYRAETPLISNOGPEIMDKLYDKFTHIVSDJPAKL 341  
DB 329 KSPMAALPVLVGFKKWVILKMKI-----SATAPELKNVKEHMF-KIVYVKKKKK 340

RESULT 10  
US-08-540-804-14  
Sequence 14; Application US/08540804  
Patent No. 5919666  
GENERAL INFORMATION:  
APPLICANT: Koleske, Richard A.  
APPLICANT: Koleske, Anthony J.



QY 267 IV-----RIGSFT--IDHTEGDLDSV-----E 288  
DB 303 AIDFIDJIKOTAAVYRPGSSVLSNAGTINPLSSDGAATFOGHHKNNSSSOPASVHS 362  
QY 289 MGFNKKWRCF---KTKVVAFTFETISNQHSPFIM-----DKLY-LKHHLIVS 335  
DB 363 VQFNKTSKDFSPNFTGTHKMLSP--SDGFAPALNTPNNINENELFNPKQTTYSN 420  
QY 336 DLEAKLPKT 344  
DB 421 DLENSPLKT 429

RESULT 12  
US-08-521-872-14  
Sequence 14, Application US/08/521872  
Patent No. 6015682  
GENERAL INFORMATION:  
APPLICANT: Young, Richard A.  
APPLICANT: Koleske, Anthony J.  
APPLICANT: Thompson, Craig M.  
APPLICANT: Chao, David M.  
TITLE OF INVENTION: No. 6015682 Factors Which Modify Gene  
TITLE OF INVENTION: Transcription and Methods of Use Therefor  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hamilton, Reay, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/521,872  
FILING DATE: 31-AUG-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/218,265  
FILING DATE: 25-MAR-1994  
AUTOPNRY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/CKET NUMBER: WH194-03A3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-9540  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1420 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-521-872-14

Query Match 48, Score 79, Pos 3, Length 1420,  
Best Local Similarity 20.08, P-Val 41,  
Matches 86, Conservative 56, Mismatches 161, Indels 126, Gaps 18.

QY 10 SYAONSSTQGVASMAQPALENAFTLFSPKTHGALN----- 47  
DB 33 SKAONDWSIWMELKRVQKTLVALSKRWLFSTINDWVLPALPRKVPSPKRCN 62  
QY 48 -AALDGCAGNTPAVTSLKRMKCKGRLNVTLELVYMLDLFNFNLFKGLSSE 106  
DB 93 FTAYSKRNLPPIHYA-----LFLKALRKRYINLALGS-HNKLIO----- 131

QY 107 VIGNKC---EEVRYVWGV-PGSP-----HGRLPPRNSLHVHSSVYHWITQ 150  
DB 132 -FVNACISI SGVRYVYVQLEPHLPVW-DLTVSLQAKMNCIYVMKRFENLFESLSKHALYL 160  
QY 151 AKGLTSPECIALNKCKIYVTSKTPVREAVYSCFHEFTMTI---NAPSGEVVWVC 206  
DB 191 APSGTRMHAPASKGTL---TTPKHTFLITTI SVSHGTINLZKRNK KVAAYVPP-- 244  
QY 207 MYILPGRVSPDSWQSCFTWEI LAMAIAYVSGCI IDPKLITFNIPSTFASLEEKD 266  
DB 245 -LGLNHTFPTIASYLPLEAKKLWPLHLIFQAVADIEN-STSGDSEPHLODALD 302  
QY 267 IV-----RIGSFT--IDHTEGDLDSV-----E 288  
DB 303 AIDFIDJIKOTAAVYRPGSSVLSNAGTINPLSSDGAATFOGHHKNNSSSOPASVHS 362  
QY 289 MGFNKKWRCF---KTKVVAFTFETISNQHSPFIM-----DKLY-LKHHLIVS 335  
DB 363 VQFNKTSKDFSPNFTGTHKMLSP--SDGFAPALNTPNNINENELFNPKQTTYSN 420  
QY 336 DLEAKLPKT 344  
DB 421 DLENSPLKT 429

RESULT 13  
US-08-590-399-14  
Sequence 14, Application US/08/590399  
Patent No. 6214588  
GENERAL INFORMATION:  
APPLICANT: Young, Richard A.  
APPLICANT: Koleske, Anthony J.  
APPLICANT: Thompson, Craig M.  
APPLICANT: Chao, David M.  
TITLE OF INVENTION: No. 6214588 Factors Which Modify Gene  
TITLE OF INVENTION: Transcription and Methods of Use Therefor  
NUMBER OF SEQUENCES: 39  
CORRESPONDENCE ADDRESS:  
ADDRESS: Hamilton, Reay, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
CITY: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/590,399  
FILING DATE: 26-JAN-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/540,804  
FILING DATE: 11-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/521,872  
FILING DATE: 31-AUG-1995  
AUTOPNRY/AGENT INFORMATION:  
NAME: Granahan, Patricia  
REGISTRATION NUMBER: 32,227  
REFERENCE/CKET NUMBER: WH194-03A3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-861-9540  
TELEFAX: 617-861-9540  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1420 amino acids  
TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US 08-590-899-14

Query Match: 4.3%, Score 79, DB 4, Length 1420,  
Best Local Similarity: 28.0%, Pred. No. 41;  
Matches: 86; Conservative: 56; Mismatches: 161; Indels: 126; Gaps: 18.

UY 10 SYAONSSFTQVAVSMAOPALENNAVELFESRDHLOALN----- 47  
DB 33 SKAONPMSTQMEPMERKPKTLVALISPMWESINDEPVPFVLEHKVSPDKIGT 92  
UY 48 -AADIQCAAPNTEFAVISTIKRMKEKCELOCOILEGVINDLEFGDNFNTLFGLSSE 106  
DB 93 FTADYSKPNLPPIRYA-----LPIKALRKITYINLAIGS HNKLTG----- 131  
UY 107 VIGNKC---EVLNCCYMGV- PGSE-----HGRLFPNLSLHVSSSYVHWLTQ 150  
DB 132 -FQMACISLGGVNYLVQLEPHLFVWQEDTVSLCAKNMGLVPMKEFNEESELKHALYL 190  
UY 151 APKQTTREBTAIINKKITYISKISFVVPFATISCHRETFMFL-----NAPSELEVVPAGC 206  
DB 191 APSTIKRMHAPASKQGYL-----ITPKHTELLTLTSLVSHGINLQKNMKVAVVPO-- 244  
UY 207 MVLLRQKQSTPSFMSQFTWELLAMALVSGTILPECKLTFNLSPYASLEFVKD 265  
DB 245 LGHLNHTTITIASYLTLLTAKKILVWPLHLIFAGVVALEN SLSISQSTPHLQALD 302  
UY 267 IVE-----PSTST--THLEGLDLSV-----E 288  
DB 303 AIDDFIQKOTAVYKTPGSSVGLSSNIASTNPLSSGAVTEQFQHYKKNSSISQPASVHS 362  
UY 289 MOENDKAVRGE---KTKVVRATFTEPISNOGPELM-----DKLY-DKTFHIYVS 335  
DB 363 VOENNKISPEKSPNTGIDKMLSP--SDQFAPAFINPNNNINBNELNDKQTVSN 420  
UY 336 DEAKLPKT 344  
DB 421 DEENSPKLT 429

RESULT 14  
US-08-286-819A-27  
Sequence 27, Application US-08-86019A  
Patent No. 5871910  
GENERAL INFORMATION:  
APPLICANT: ARTHUR, MICHEL  
APPLICANT: PIKTA, MALEN, SYLVIE  
APPLICANT: MOLINAS, CATHERINE  
TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE  
TITLE OF INVENTION: EXPRESSION OF PESTICANT TO POLYPEPTIDES, IN PARTICULAR  
TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR  
NUMBER OF SEQUENCES: 54  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSJADT,  
ADDRESSEE: P.C.  
STREET: 1765 S. Jefferson Boyle Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US-08/286,819A  
FILING DATE: 05-AUG-1994  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/174,682  
FILING DATE: 28-DEC-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/917,146  
FILING DATE: 10-AUG-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/FR/91/00855  
FILING DATE: 29-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 9013579  
FILING DATE: 31-OCT-1990  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: OBLON, NO. 5871910man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/WORKER NUMBER: 000-000-0 PCT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-4000  
TELEFAX: (703) 413-4220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2296 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-286-819A-27

Query Match: 4.3%, Score 78.5, DB 2, Length 2296;  
Best Local Similarity: 18.0%, Pred. No. 99;  
Matches: 79; Conservative: 68; Mismatches: 103; Indels: 147; Gaps: 22;

UY 53 CAAGPNTFAVIST-----IKRME-----KKC---RELN---CUTDEL----- 84  
DB 241 GQDPRESFKTISAVDRDNGYIKRPFQCGKIASNFKCTIYPMKTSFMLQTIESLVH 300  
UY 85 QVYINDLPQNDFTLFGKLSSEV---IGKCKEVCYMGVPGSGHGLFPNLSLHVHS 141  
DB 301 KYLEN-----SLTYEKRGVNR-----IHGLYGR-----IHT 340  
UY 142 SYVIMHLLGAKCLISEPRL-----ALR---FCKIYLSLSFVVEAVLSQEHDEPT 191  
DB 331 ANSLNLY-----LTSNELLPTNNVKGILNMLKKKSTVAPDSILKTFQIMPKLEYKRN 385  
UY 192 MLN-----APSEVFNQGMVLLKGR---GQSDVPMQSTTW 228  
DB 386 MYNQCEFTTNVSPALYKISEV--NNFQCLPMGNLFKEKELIKYQPSMAEPPDK 444  
UY 429 ELIAMA-----LALVSGQLIDEDKIDPENTISYFASG-----EYKQIVE R 270  
DB 444 KPTILLNSSFMLMLNSAL--SPKLLPANICEVYTSQKLLVDEBHHADVLELYK 501  
UY 271 DQSEFT-----DRIEGFYDSVIMQENQVWPCEKTYVVPATFPIISNOGPELM 323  
DB 502 NENYIVFYKYAKFALECIDKSEIDL-----ALIDIMDTSGITG 544  
UY 324 KLYDKTH--IYVSDLEAKLPRTSILL 349  
DB 545 KIPDKHTYPLMKGIDYKVGKITGLTI 572

RESULT 15  
US-08-980-357-27  
Sequence 27, Application US-08980357  
Patent No. 6013508  
GENERAL INFORMATION:  
APPLICANT: ARTHUR, MICHEL  
APPLICANT: PIKTA, MALEN, SYLVIE  
APPLICANT: MOLINAS, CATHERINE

APPLICANT: COBRALIN, PATRICE  
 TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE  
 TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPOLYMER  
 TITLE OF INVENTION: IN GRAM POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE  
 TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS  
 NUMBER OF SEQUENCES: 54  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: GILSON, SALVATOR, MONTPELLIER, MAIRIE & NEUSAILL  
 STREET: 1755 S. Jefferson Davis Highway, Suite 400  
 CITY: Arlington  
 STATE: Virginia  
 COUNTRY: U.S.A.  
 ZIP: 22202  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/980,457  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/286,819  
 FILING DATE: 05-AUG-1994  
 APPLICATION NUMBER: US 08/174,682  
 FILING DATE: 28-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/917,146  
 FILING DATE: 10-AUG-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCT/FR/91/00855  
 FILING DATE: 29-OCT-1991  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: FR 9013579  
 FILING DATE: 31-OCT-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OBIOL, No. 6013508man P.  
 REGISTRATION NUMBER: 24,618  
 REFERENCE/DOCKET NUMBER: 660-060-0 PCT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (703) 413-3000  
 TELEFAX: (703) 413-2220  
 TELEX: 248855 OPAT UR  
 INFORMATION FOR SEQ ID NO: 27:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2296 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-980-357-27

Query Match 4 % Score 78 5; PR 3; length 2296;  
 Best local similarity 18.0% Pred. No. 99;  
 Matches 70; Conservative 68; Mismatches 103; Indels 147; Gaps 22;

QY 53 CAAGPNTFAVIST-----IKRME-----KRC---RELN---COTLEL----- 84  
 Db 241 CPTPEPFSFKTISAVEKIPKNGYVYIKFRQDEQRIASNEKKCIYKKMISPMLOITISLVVH 300  
 QY 85 QVYINDLFGNDPNTLFGKLSSEV---IGNKCEVPCYVMGVPQSGFGRLEPRNSLTVHS 141  
 Db 301 KIYLN-----SITYEIKROYNKK-----IHGLIYQR-----IHT 330  
 QY 142 SYSVHMLTQAPKGLTSEGL-----ALN---GKTIYISKISPPVVRKAYLSQFHEDFT 191  
 Db 331 ANSLMLV-----LTNSEILFGDNVKGKLMRLKRESIKVDRSIIKITQELMRKLYKEGN 385  
 QY 192 MFLN-----ARSGFVVPNGCMVLLPGR-----QSNPSMDSQFTW 228  
 Db 386 MIVMOIGETINVSASLYRKLSEV--NNPFCIPLMGNIFKEKETIKYQPPSDAEKPFDK 443

QY 229 ELIAMA - LAELNSQGLTEADLITNIESYASL-----EYKIVE---P 270  
 Db 444 KELLINSSSEIMMLINSAL--SKETIPANIVTEVYIMSEKIIIMVCHETIAIWEYELK 501  
 QY 271 YASFTI-----PIHCPPTLSVEMQPNPKWPGCKPTKPVKATFTTITSNQSPFTMD 323  
 Db 502 NENTIVHYIYIAEALGCTIKNSIDE-----ALIDIMLPQTSGLTIQ 544  
 QY 424 KLYDKFTH--IWSDBKAKLPKITSIL 349  
 Db 545 KIRDKHTYPIIMLGKDTEDKITGLTI 572

Search completed: August 31, 2001, 15:00:28  
 Job time: 2161 sec

Sat Sep 1 11:45:35 2001

us-09-577-657a-1.rai

---

Page 10



```

XX This invention provides a novel DNA molecule (1) encoding N-methyl
CC transferase (N-met) with the enzyme activities of 7-methyl xanthine N3
CC methyl transferase, theobromine N1 methyl transferase and paraxanthine
CC N3 methyl transferase. The DNA or RNA encoding the N-met is useful for
CC the enhancement of caffeine production in microorganisms or plants, and
CC the encoded N-met is also used for caffeine synthesis. The methods and
CC the DNA may also be used for producing N-met that can be used as an
CC enzyme for industrial, food or medical use, e.g. for producing compounds
CC related to caffeine metabolism by modifying caffeine biosynthesis and
CC metabolism of caffeine productive plants, plant tissues or cells and for
CC modifying caffeine biosynthesis and metabolism of caffeine productive
CC plants, plant tissues or cells, therefore modifying the production rate
CC of caffeine metabolism related compounds.
XX
SU Sequence 356 AA:

Query Match 100.0%; Score 1847; E-Value 2.2; Length 356;
Best Local Similarity 100.0%; Field No. 2,96-177;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

UY 1 FNNKGESESYAONSSFTQVAVASMAQALBNVETLFESDHFHUALNADUAGAPNIF 60
    |||||||
DB 1 tmrpgessyqsssttqvaamaqalbnvettlfdhpaluadlgyagpft 60

UY 61 AVISTEKRMKEKKKPEI NQTHLHGVYNIENFENENILEKGLSESVASNGCEVPCVYM 120
    |||||||
DB 61 dvistekrmekkkropeitnqthlhgvynienfeneenilekglseesvasngcevpvyym 120

UY 121 GVSQSTFEGHLEPPNSI HLHSSYSVHMLTQAKGTTSPGLAINCKYISKTPPVVVF 180
    |||||||
DB 121 gvsptfgrltpnsihlvhssysvmltqakgltspglainckyisktpvvvfe 180

UY 181 AVI SQEPFETMELNMRSGVYVNRVMVLI RPPQCSFNSQMSQSTFWILLAMAFAELVS 240
    |||||||
DB 181 ayslqfethetmlnarsqvynrcvmvllrppqcsfnsqmsqstfwllamafaelvs 240

UY 241 QGLDDEKDKJTFNIPSYFASLEKVDIVRQPSFTTHIESFPLASVEMENKMWPGSK 300
    |||||||
DB 241 qglldkdyldtrpsyfasleekvdivrqpstthiesfplasvemenkmpgsk 300

UY 301 FTKVVAATFPPLISNGFGRPMKLYDKFTHIVASPLEKELKRTISLIVLSKIDG 356
    |||||||
DB 301 ftkvvaatfpplisngfgrpmklydkftthivaspleslekrlkrtislivlskidg 356

RESULT 2
AAI22190
ID AAG22190 standard; protein; 386 AA.
XX
AC AAG22190:
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25024
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN PP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-SEP-2000; 2000EP-0301439.
XX
PR 25-SEP-1999; 99US-0121825
PR 05-MAY-1999; 99US-0123180.
PR 09-MAY-1999; 99US-0123548.
PR 21-MAY-1999; 99US-0125788.

```

```

PR 25-MAY-1999; 99US-0126364.
PR 23-MAY-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128334.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129445.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130810.
PR 25-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138440.
PR 10-JUN-1999; 99US-0138447.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139849.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140951.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142056.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.

```





XX protein identification: signal transduction pathway: metabolic pathway:  
 KM hybridisation assay: genetic mapping, gene expression control, promoter:  
 KM termination sequence  
 XX  
 XX Arabidopsis thaliana.  
 EN EP1034405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000: 2000EP-0401439.  
 XX  
 PR 25-FEB-1999: 99US-0121825.  
 PR 05-MAR-1999: 99US-0123180.  
 PR 09-MAR-1999: 99US-0123548.  
 PR 23-MAR-1999: 99US-0125788.  
 PR 28-MAR-1999: 99US-0126264.  
 PR 29-MAR-1999: 99US-0126785.  
 PR 01-APR-1999: 99US-0127462.  
 PR 08-APR-1999: 99US-0128234.  
 PR 08-APR-1999: 99US-0128714.  
 PR 16-APR-1999: 99US-0129845.  
 PR 19-APR-1999: 99US-0130077.  
 PR 21-APR-1999: 99US-0130449.  
 PR 21-APR-1999: 99US-0130510.  
 PR 21-APR-1999: 99US-0130891.  
 PR 28-APR-1999: 99US-0131449.  
 PR 40-APR-1999: 99US-0132048.  
 PR 40-APR-1999: 99US-0132407.  
 PR 04-MAY-1999: 99US-0132484.  
 PR 05-MAY-1999: 99US-0132486.  
 PR 06-MAY-1999: 99US-0132486.  
 PR 06-MAY-1999: 99US-0132487.  
 PR 07-MAY-1999: 99US-0132863.  
 PR 11-MAY-1999: 99US-0134256.  
 PR 14-MAY-1999: 99US-0134218.  
 PR 14-MAY-1999: 99US-0134219.  
 PR 14-MAY-1999: 99US-0134231.  
 PR 18-MAY-1999: 99US-0134370.  
 PR 18-MAY-1999: 99US-0134768.  
 PR 19-MAY-1999: 99US-0134941.  
 PR 20-MAY-1999: 99US-0135124.  
 PR 21-MAY-1999: 99US-0135124.  
 PR 24-MAY-1999: 99US-0135353.  
 PR 25-MAY-1999: 99US-0136021.  
 PR 27-MAY-1999: 99US-0136392.  
 PR 28-MAY-1999: 99US-0136782.  
 PR 01-JUN-1999: 99US-0137222.  
 PR 03-JUN-1999: 99US-0137528.  
 PR 04-JUN-1999: 99US-0137502.  
 PR 07-JUN-1999: 99US-0137724.  
 PR 08-JUN-1999: 99US-0138094.  
 PR 10-JUN-1999: 99US-0138540.  
 PR 10-JUN-1999: 99US-0138847.  
 PR 14-JUN-1999: 99US-0139119.  
 PR 16-JUN-1999: 99US-0139452.  
 PR 16-JUN-1999: 99US-0139453.  
 PR 17-JUN-1999: 99US-0139492.  
 PR 18-JUN-1999: 99US-0139454.  
 PR 18-JUN-1999: 99US-0139460.  
 PR 18-JUN-1999: 99US-0139461.  
 PR 18-JUN-1999: 99US-0139462.  
 PR 18-JUN-1999: 99US-0139463.  
 PR 18-JUN-1999: 99US-0139750.  
 PR 21-JUN-1999: 99US-0139763.  
 PR 22-JUN-1999: 99US-0139817.  
 PR 22-JUN-1999: 99US-0139899.

---

PR 23-JUN-1999: 99US-0140353.  
 PR 23-JUN-1999: 99US-0140354.  
 PR 24-JUN-1999: 99US-0140604.  
 PR 28-JUN-1999: 99US-0140873.  
 PR 29-JUN-1999: 99US-0140941.  
 PR 30-JUN-1999: 99US-0141287.  
 PR 01-JUL-1999: 99US-0141842.  
 PR 01-JUL-1999: 99US-0142154.  
 PR 02-JUL-1999: 99US-0142055.  
 PR 06-JUL-1999: 99US-0142390.  
 PR 08-JUL-1999: 99US-0142803.  
 PR 09-JUL-1999: 99US-0142920.  
 PR 12-JUL-1999: 99US-0142977.  
 PR 13-JUL-1999: 99US-0143542.  
 PR 14-JUL-1999: 99US-0143624.  
 PR 15-JUL-1999: 99US-0144005.  
 PR 16-JUL-1999: 99US-0144085.  
 PR 16-JUL-1999: 99US-0144086.  
 PR 19-JUL-1999: 99US-0144325.  
 PR 19-JUL-1999: 99US-0144331.  
 PR 19-JUL-1999: 99US-0144332.  
 PR 19-JUL-1999: 99US-0144333.  
 PR 19-JUL-1999: 99US-0144334.  
 PR 19-JUL-1999: 99US-0144335.  
 PR 20-JUL-1999: 99US-0144352.  
 PR 20-JUL-1999: 99US-0144632.  
 PR 21-JUL-1999: 99US-0144884.  
 PR 21-JUL-1999: 99US-0144814.  
 PR 21-JUL-1999: 99US-0145086.  
 PR 21-JUL-1999: 99US-0145088.  
 PR 22-JUL-1999: 99US-0145085.  
 PR 22-JUL-1999: 99US-0145087.  
 PR 22-JUL-1999: 99US-0145087.  
 PR 22-JUL-1999: 99US-0145089.  
 PR 22-JUL-1999: 99US-0145192.  
 PR 23-JUL-1999: 99US-0145145.  
 PR 23-JUL-1999: 99US-0145218.  
 PR 23-JUL-1999: 99US-0145224.  
 PR 26-JUL-1999: 99US-0145276.  
 PR 27-JUL-1999: 99US-0145913.  
 PR 27-JUL-1999: 99US-0145919.  
 PR 27-JUL-1999: 99US-0145919.  
 PR 28-JUL-1999: 99US-0145951.  
 PR 02-AUG-1999: 99US-0146386.  
 PR 02-AUG-1999: 99US-0146388.  
 PR 02-AUG-1999: 99US-0146389.  
 PR 03-AUG-1999: 99US-0147034.  
 PR 04-AUG-1999: 99US-0147204.  
 PR 04-AUG-1999: 99US-0147302.  
 PR 05-AUG-1999: 99US-0147392.  
 PR 05-AUG-1999: 99US-0147260.  
 PR 06-AUG-1999: 99US-0147303.  
 PR 06-AUG-1999: 99US-0147416.  
 PR 09-AUG-1999: 99US-0147494.  
 PR 09-AUG-1999: 99US-0147935.  
 PR 10-AUG-1999: 99US-0148171.  
 PR 11-AUG-1999: 99US-0148319.  
 PR 12-AUG-1999: 99US-0148341.  
 PR 13-AUG-1999: 99US-0148565.  
 PR 13-AUG-1999: 99US-0148684.  
 PR 16-AUG-1999: 99US-0148688.  
 PR 17-AUG-1999: 99US-0149175.  
 PR 18-AUG-1999: 99US-0149426.  
 PR 20-AUG-1999: 99US-0149722.  
 PR 20-AUG-1999: 99US-0149723.  
 PR 20-AUG-1999: 99US-0149929.  
 PR 23-AUG-1999: 99US-0149932.  
 PR 23-AUG-1999: 99US-0149930.  
 PR 25-AUG-1999: 99US-0150566.  
 PR 26-AUG-1999: 99US-0150884.  
 PR 27-AUG-1999: 99US-0151065.  
 PR 27-AUG-1999: 99US-0151066.  
 PR 27-AUG-1999: 99US-0151080.  
 PR 30-AUG-1999: 99US-0151303.

```

PR 31-SEP-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151430.
PR 07-SEP-1999; 9905-0152463.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155446.
PR 24-SEP-1999; 9905-0155659.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156595.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157855.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160940.
PR 22-OCT-1999; 9905-0160941.
PR 22-OCT-1999; 9905-0160949.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161359.
PR 26-OCT-1999; 9905-0161360.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161902.
PR 28-OCT-1999; 9905-0161903.
PR 29-OCT-1999; 9905-0162142.

```

Query Match 38.84; Score 716.5; PR 21; Length 389;  
 Best Local Similarity 39.34; Pred No 16e-64;  
 Matches 148; Conservative 72; Mismatches 130; Indels 27; Gaps 5;

```

OY 2 MNRGSSVYKQNSSTLYGVASMAQPALENAVERIPSKRPHDAIIMALIGDCAQINPIA 61
DB 9 MUKGUELSYKAKSLASALISLIRVMDALKKIMNSSEISIGIADIGSSQPSNLI 68
OY 62 VISTIKRMKKEKKEIENQITLQVYINDLEGNPNTLEFKGLSS--EVIGNKKFEV---- 115
DB 69 slmshvdlhlnpdlaprpelvtslndlsrdlnyiaslsrlefydymmkkglygr 128
OY 116 ----EYVMVPSGFGHGLPPKNSLHLVHSSYVHWLILAPKDLISPELAL-----NKA 169
DB 129 gggscfvsavyslygrlfrtslthvssslhwlslypcreaeekedltladlenng 188
OY 167 KIVSKTSPPVVRKAVYLSGHEHETMFLNAPSCVVVINGMVLILKQKESDPSUMNSCH 246
DB 189 klysktsprkshakayalqptdlvlltrseelvpqgmvislqtrslqpttosecy 248
OY 227 TWELLAMIAELVQCLIDELDKLITFNIFSYFASLEVKDILVKDQSEFTLDHIEGFLD- 285
DB 249 qwellagymmakgyleekidafnapyaasaelkwnlckegsfslldrlslpdlw 308
OY 286 ---SVEMDENRKNV-----GEKFTKVVKATLEPLISNUGPELIMKLYDKTHIVV 344

```

```

DB 309 eggslseesypalatrskpealaeqrwslrtaycpmleptlgenwmdelleryakng 368
OY 335 SDLEAKLPKTSITILV 351
DB 369 cyfvsprpvalvls 385

RESULT 4
AAC22188
ID AAC22188 standard; Protein; 405 AA.
AC XX
NC AAC22188.
XX XX
XX 17-OCT-2000 (first entry)
XX XX
XX Arabidopsis thaliana protein fragment, SEQ ID NO: 25022.
XX XX
XX Protein identification: signal transduction pathway; metabolic pathway;
XX XX hybridisation assay; genetic mapping; gene expression control; promoter;
XX XX termination sequence.
XX XX Arabidopsis thaliana.
XX XX EP1033405-A2.
XX XX
XX 05-SEP-2000.
XX XX
XX 24-FEB-2000; 2000EP-0301439.
XX XX
XX 25-FEB-1999; 9905-0121825.
XX XX 05-MAR-1999; 9905-0123180.
XX XX 09-MAR-1999; 9905-0123548.
XX XX 23-MAR-1999; 9905-0125788.
XX XX 25-MAR-1999; 9905-0126264.
XX XX 29-MAR-1999; 9905-0126785.
XX XX 02-APR-1999; 9905-0127452.
XX XX 06-APR-1999; 9905-0128234.
XX XX 08-APR-1999; 9905-0128714.
XX XX 16-APR-1999; 9905-0129845.
XX XX 19-APR-1999; 9905-0130077.
XX XX 21-APR-1999; 9905-0130449.
XX XX 23-APR-1999; 9905-0130510.
XX XX 23-APR-1999; 9905-0130891.
XX XX 28-APR-1999; 9905-0131449.
XX XX 30-APR-1999; 9905-0132048.
XX XX 30-APR-1999; 9905-0132407.
XX XX 04-MAY-1999; 9905-0132484.
XX XX 05-MAY-1999; 9905-0132485.
XX XX 06-MAY-1999; 9905-0132486.
XX XX 06-MAY-1999; 9905-0132487.
XX XX 07-MAY-1999; 9905-0132863.
XX XX 11-MAY-1999; 9905-0134256.
XX XX 14-MAY-1999; 9905-0134218.
XX XX 14-MAY-1999; 9905-0134219.
XX XX 14-MAY-1999; 9905-0134221.
XX XX 14-MAY-1999; 9905-0134370.
XX XX 18-MAY-1999; 9905-0134768.
XX XX 19-MAY-1999; 9905-0134941.
XX XX 20-MAY-1999; 9905-0135124.
XX XX 21-MAY-1999; 9905-0135353.
XX XX 24-MAY-1999; 9905-0135629.
XX XX 25-MAY-1999; 9905-0136021.
XX XX 27-MAY-1999; 9905-0136392.
XX XX 28-MAY-1999; 9905-0136782.
XX XX 01-JUN-1999; 9905-0137252.
XX XX 03-JUN-1999; 9905-0137528.
XX XX 04-JUN-1999; 9905-0137502.
XX XX 07-JUN-1999; 9905-0137724.
XX XX 08-JUN-1999; 9905-0138094.
XX XX 10-JUN-1999; 9905-0138540.
XX XX 14-JUN-1999; 9905-0138847.
XX XX 14-JUN-1999; 9905-0139119.

```

```

PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0139753.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139819.
PR 23-JUN-1999; 990S-0140353.
PR 24-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142300.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143422.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144085.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144425.
PR 19-JUL-1999; 990S-0144431.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 20-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144352.
PR 20-JUL-1999; 990S-0144632.
PR 21-JUL-1999; 990S-0144884.
PR 21-JUL-1999; 990S-0144814.
PR 21-JUL-1999; 990S-0145086.
PR 22-JUL-1999; 990S-0145088.
PR 22-JUL-1999; 990S-0145085.
PR 22-JUL-1999; 990S-0145087.
PR 22-JUL-1999; 990S-0145089.
PR 22-JUL-1999; 990S-0145192.
PR 23-JUL-1999; 990S-0145145.
PR 23-JUL-1999; 990S-0145218.
PR 23-JUL-1999; 990S-0145224.
PR 25-JUL-1999; 990S-0145276.
PR 27-JUL-1999; 990S-0145913.
PR 27-JUL-1999; 990S-0145918.
PR 27-JUL-1999; 990S-0145919.
PR 28-JUL-1999; 990S-0145951.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147038.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.

```

```

PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148365.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149722.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149930.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150884.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151438.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0156559.
PR 28-SEP-1999; 990S-0156458.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 08-OCT-1999; 990S-0158369.
PR 12-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159331.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159638.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160770.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 23-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.
PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

```

Query Match 38.8%; Score 716.5; DB 21; Length 405;

Best Local Similarity 39.8%; Prof No 17e+63;

Matches 149; Conservative 72; Mismatches 109; Indels 27; Gaps 5;

```

QY 2 MNPGEPSYVQNSPTQGVASWAGPATFNAPVTFPSRPHIQAIAANVGGACAGPTFA 61
  |||||
2b 25 mdkqgqetcyahstqstqstlslfrvcdaklhlmshtvststgjdltqstpsl 84

```

```

OY 62 VISTIKMKKKKREINQUTLHGVYINLPNINENTLPRKGLSS--EVYGNCKEEV---- 115
DB 85 sislivdtlnldpdlidrpwplrvslndlpndfioylcaslpefydyvnmkglqfgr 144
OY 116 ----PYVVMVWVQSPHKLPLKNSLHVSSVSHVHITQAPKGLSKRELAL-----MKG 166
DB 145 ggesctsvsvpsfgygrlfrtrstlhfvsesslholseqprroackgrttrtdlcmng 204
OY 167 KIYIKSTSPVPEAVLSPHPLHPLHPLMAPSLQVVPVNRKQWLLHPLQWNPSPMNSP 226
DB 265 qwellagallsmakqglreekldatnaryaaaseelkmvlekegsislrlrslpdlw 324
OY 227 TWETLMANAFVSGNIDELKIDFNIQSYTASLHRYKQVLEKNSPTLHPLHPLHPL 285
DB 265 qwellagallsmakqglreekldatnaryaaaseelkmvlekegsislrlrslpdlw 324
OY 286 ---SVEQNDKKVNR-----GERTKVVAFTEPLISNQGPEIMDKYDEKTHIV 334
DB 325 egsylseesydlatrskpealagrvsnlllavvegmleprlgevmdeleterfakivg 384
OY 335 SDLEAKLPKTTSTITIV 351
DB 385 eylvsspryaivylsl 401

FESTI 5
ID AAC12687 standard: Protein: 348 AA
AC AAC12687.
XX 17-OCT-2000 (first entry)
DB Arabidopsis thaliana protein fragment SPY II No. 1194
XX
XX protein identification, signal transduction pathway, metabolic pathway,
KM hybridisation assay, genetic mapping, gene expression control, promoter,
XX termination sequence.
XX Arabidopsis thaliana.
XX OS
XX EP103405-A2
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000-2000EP-0101439
XX
XX 25-FEB-1999: 990S-0121825.
XX 05-MAY-1999: 990S-0123180.
XX 09-MAY-1999: 990S-0123548.
XX 23-MAY-1999: 990S-0125788.
XX 25-MAY-1999: 990S-0126785.
XX 29-MAY-1999: 990S-0127462.
XX 01-APR-1999: 990S-0128254.
XX 08-APR-1999: 990S-0128714.
XX 16-APR-1999: 990S-0129845.
XX 19-APR-1999: 990S-0130077.
XX 21-APR-1999: 990S-0130449.
XX 23-APR-1999: 990S-0130510.
XX 24-APR-1999: 990S-0130891.
XX 28-APR-1999: 990S-0131449.
XX 30-APR-1999: 990S-0132048.
XX 04-MAY-1999: 990S-0132407.
XX 05-MAY-1999: 990S-0132484.
XX 06-MAY-1999: 990S-0132486.
XX 07-MAY-1999: 990S-0132863.
XX 11-MAY-1999: 990S-0134256.
XX 14-MAY-1999: 990S-0134318.
XX 14-MAY-1999: 990S-0134319.
XX 14-MAY-1999: 990S-0134321.

```

```

PR 14-MAY-1999: 990S-0134370.
PR 18-MAY-1999: 990S-0134768.
PR 19-MAY-1999: 990S-0134941.
PR 20-MAY-1999: 990S-0135124.
PR 21-MAY-1999: 990S-0135353.
PR 24-MAY-1999: 990S-0135629.
PR 25-MAY-1999: 990S-0136021.
PR 27-MAY-1999: 990S-0136392.
PR 28-MAY-1999: 990S-0136782.
PR 01-JUN-1999: 990S-0137032.
PR 03-JUN-1999: 990S-0137528.
PR 04-JUN-1999: 990S-0137502.
PR 07-JUN-1999: 990S-0137724.
PR 08-JUN-1999: 990S-0138094.
PR 10-JUN-1999: 990S-0138401.
PR 10-JUN-1999: 990S-0138847.
PR 14-JUN-1999: 990S-0139119.
PR 16-JUN-1999: 990S-0139452.
PR 16-JUN-1999: 990S-0139453.
PR 17-JUN-1999: 990S-0139702.
PR 18-JUN-1999: 990S-0139454.
PR 18-JUN-1999: 990S-0139456.
PR 18-JUN-1999: 990S-0139457.
PR 18-JUN-1999: 990S-0139458.
PR 18-JUN-1999: 990S-0139459.
PR 18-JUN-1999: 990S-0139460.
PR 18-JUN-1999: 990S-0139461.
PR 18-JUN-1999: 990S-0139462.
PR 18-JUN-1999: 990S-0139463.
PR 18-JUN-1999: 990S-0139750.
PR 18-JUN-1999: 990S-0139763.
PR 21-JUN-1999: 990S-0139817.
PR 22-JUN-1999: 990S-0139804.
PR 23-JUN-1999: 990S-0140353.
PR 23-JUN-1999: 990S-0140354.
PR 24-JUN-1999: 990S-0140695.
PR 28-JUN-1999: 990S-0140823.
PR 29-JUN-1999: 990S-0140991.
PR 30-JUN-1999: 990S-0141287.
PR 01-JUL-1999: 990S-0141842.
PR 01-JUL-1999: 990S-0142154.
PR 02-JUL-1999: 990S-0142055.
PR 06-JUL-1999: 990S-0142760.
PR 08-JUL-1999: 990S-0142803.
PR 09-JUL-1999: 990S-0142820.
PR 12-JUL-1999: 990S-0142977.
PR 13-JUL-1999: 990S-0143642.
PR 14-JUL-1999: 990S-0143624.
PR 16-JUL-1999: 990S-0144005.
PR 16-JUL-1999: 990S-0144085.
PR 16-JUL-1999: 990S-0144086.
PR 19-JUL-1999: 990S-0144325.
PR 19-JUL-1999: 990S-0144331.
PR 19-JUL-1999: 990S-0144332.
PR 19-JUL-1999: 990S-0144333.
PR 19-JUL-1999: 990S-0144334.
PR 19-JUL-1999: 990S-0144335.
PR 20-JUL-1999: 990S-0144351.
PR 20-JUL-1999: 990S-0144352.
PR 20-JUL-1999: 990S-0144884.
PR 21-JUL-1999: 990S-0144814.
PR 21-JUL-1999: 990S-0145086.
PR 21-JUL-1999: 990S-0145088.
PR 22-JUL-1999: 990S-0145085.
PR 22-JUL-1999: 990S-0145087.
PR 22-JUL-1999: 990S-0145089.
PR 22-JUL-1999: 990S-0145192.
PR 23-JUL-1999: 990S-0145145.
PR 23-JUL-1999: 990S-0145148.
PR 23-JUL-1999: 990S-0145224.
PR 26-JUL-1999: 990S-0145276.
PR 27-JUL-1999: 990S-0145913.

```

```

PR 27-JUL-1999; 990S-0145618.
PR 27-JUL-1999; 990S-0145619.
PR 28-JUL-1999; 990S-0145931.
PR 02-AUG-1999; 990S-0146386.
PR 02-AUG-1999; 990S-0146388.
PR 02-AUG-1999; 990S-0146389.
PR 03-AUG-1999; 990S-0147048.
PR 04-AUG-1999; 990S-0147204.
PR 04-AUG-1999; 990S-0147302.
PR 05-AUG-1999; 990S-0147192.
PR 05-AUG-1999; 990S-0147260.
PR 06-AUG-1999; 990S-0147303.
PR 06-AUG-1999; 990S-0147416.
PR 09-AUG-1999; 990S-0147493.
PR 09-AUG-1999; 990S-0147935.
PR 10-AUG-1999; 990S-0148171.
PR 11-AUG-1999; 990S-0148319.
PR 12-AUG-1999; 990S-0148341.
PR 13-AUG-1999; 990S-0148565.
PR 13-AUG-1999; 990S-0148684.
PR 16-AUG-1999; 990S-0149368.
PR 17-AUG-1999; 990S-0149175.
PR 18-AUG-1999; 990S-0149426.
PR 20-AUG-1999; 990S-0149723.
PR 20-AUG-1999; 990S-0149929.
PR 23-AUG-1999; 990S-0149902.
PR 23-AUG-1999; 990S-0149930.
PR 25-AUG-1999; 990S-0150566.
PR 26-AUG-1999; 990S-0150684.
PR 27-AUG-1999; 990S-0151065.
PR 27-AUG-1999; 990S-0151066.
PR 27-AUG-1999; 990S-0151080.
PR 30-AUG-1999; 990S-0151303.
PR 31-AUG-1999; 990S-0151338.
PR 01-SEP-1999; 990S-0151930.
PR 07-SEP-1999; 990S-0152363.
PR 10-SEP-1999; 990S-0153070.
PR 13-SEP-1999; 990S-0153758.
PR 15-SEP-1999; 990S-0154018.
PR 16-SEP-1999; 990S-0154039.
PR 20-SEP-1999; 990S-0154779.
PR 22-SEP-1999; 990S-0155139.
PR 23-SEP-1999; 990S-0155486.
PR 24-SEP-1999; 990S-0155659.
PR 28-SEP-1999; 990S-0156456.
PR 29-SEP-1999; 990S-0156596.
PR 04-OCT-1999; 990S-0157117.
PR 05-OCT-1999; 990S-0157753.
PR 06-OCT-1999; 990S-0157865.
PR 07-OCT-1999; 990S-0158029.
PR 08-OCT-1999; 990S-0158232.
PR 12-OCT-1999; 990S-0158364.
PR 13-OCT-1999; 990S-0159293.
PR 13-OCT-1999; 990S-0159294.
PR 13-OCT-1999; 990S-0159295.
PR 14-OCT-1999; 990S-0159329.
PR 14-OCT-1999; 990S-0159330.
PR 14-OCT-1999; 990S-0159431.
PR 14-OCT-1999; 990S-0159637.
PR 14-OCT-1999; 990S-0159648.
PR 18-OCT-1999; 990S-0159584.
PR 21-OCT-1999; 990S-0160741.
PR 21-OCT-1999; 990S-0160767.
PR 21-OCT-1999; 990S-0160768.
PR 21-OCT-1999; 990S-0160771.
PR 21-OCT-1999; 990S-0160814.
PR 21-OCT-1999; 990S-0160815.
PR 22-OCT-1999; 990S-0160980.
PR 22-OCT-1999; 990S-0160981.
PR 22-OCT-1999; 990S-0160989.
PR 25-OCT-1999; 990S-0161404.
PR 25-OCT-1999; 990S-0161405.

```

```

PR 25-OCT-1999; 990S-0161406.
PR 26-OCT-1999; 990S-0161359.
PR 26-OCT-1999; 990S-0161360.
PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161922.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162342.
Query Match 37.2% Score 687; DB 21; Length 349;
Best Local Similarity 40.2% Fred No 1 3e-60;
Matches 144, Conservative 70, Mismatches 128, Indels 16, Gaps 6;
QY 2 MNRPFSTSSAAGNSSTFLGVASMAAGAI-ENAVETLFSDFEHLGALNAINAGCAAPRTF 60
Db 1 msqgddqansysansrlqkkvksnakpvlrrlccommldlply-llkvoqlqssqgnsl 59
QY 61 AVISTKRMMKKCFEINCCOTELIYVNLNLFENDNLIKCL---SSEVIGNKEEVP 116
Db 60 lalfelintlnvleqhnokspeldccldlpendntlltkfvtlndcmldss--- 116
QY 117 CYWKGVPSTFGRLEPPNSLIHLWSSYVHWLTQAKGLTSTREGIALNKKIYISKTSPP 176
Db 117 cf79qpfstfysllfalselhlhlsygrhlwskv- chlvutnlytlssst- 170
QY 177 VVREAVYISQPHPIPTMPLNARSGFVVPNGCVYLIRGHO-GSDPSDMQSCFTWELLAMAI 235
Db 171 s3ykaylnqfkdftmlrlfseelstgrmlltltgrndlnpplrydchhkwlllsssl 230
QY 231 rdlvfyglvsoskldafnmpfygfwgylkvogstfoelnesbgfdlghygeddt 290
Db 296 VPDGKFTKVVAFTEPPISNQPPELMOKYDPTTHIVSDLEALFRTSLILVLSK 353
Db 291 cagmcsnglrvasepmllshlqgeelldclfskayhvcvqncnckctvsivsltk 348
RESULT 6
AAG37440
ID AAG37440 standard; protein; 349 AA.
XX
AC AAG37440;
XX
L1 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SRQ ID NO: 46036.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression controls; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25 FEB 2000; J060EP-0001439.
XX
PR 25-FEB-1999; 990S-01231825.
PR 05-MAR-1999; 990S-0123180.
PR 09-MAR-1999; 990S-0123548.
PR 23-MAR-1999; 990S-0125788.
PR 25-MAR-1999; 990S-0125784.
PR 29-MAR-1999; 990S-0125785.
PR 01-APR-1999; 990S-0127462.
PR 06-APR-1999; 990S-0128234.
PR 08-APR-1999; 990S-0128714.
PR 16-APR-1999; 990S-0129845.
PR 19-APR-1999; 990S-0130077.
PR 21-APR-1999; 990S-0130449.
PR 23-APR-1999; 990S-0130510.

```

PR 23-APR-1999: 99US-0130891.  
PR 28-APR-1999: 99US-0131449.  
PR 30-APR-1999: 99US-0132648.  
PR 04-MAY-1999: 99US-0132407.  
PR 05-MAY-1999: 99US-0132484.  
PR 06-MAY-1999: 99US-0132485.  
PR 06-MAY-1999: 99US-0132486.  
PR 07-MAY-1999: 99US-0132487.  
PR 11-MAY-1999: 99US-0132863.  
PR 14-MAY-1999: 99US-0134256.  
PR 14-MAY-1999: 99US-0134218.  
PR 14-MAY-1999: 99US-0134219.  
PR 14-MAY-1999: 99US-0134221.  
PR 18-MAY-1999: 99US-0134370.  
PR 19-MAY-1999: 99US-0134768.  
PR 19-MAY-1999: 99US-0134841.  
PR 23-MAY-1999: 99US-0135124.  
PR 23-MAY-1999: 99US-0135629.  
PR 25-MAY-1999: 99US-0136021.  
PR 27-MAY-1999: 99US-0136992.  
PR 28-MAY-1999: 99US-0136992.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.  
PR 08-JUN-1999: 99US-0138094.  
PR 10-JUN-1999: 99US-0138540.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139492.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139464.  
PR 18-JUN-1999: 99US-0139465.  
PR 21-JUN-1999: 99US-0139467.  
PR 22-JUN-1999: 99US-0139899.  
PR 23-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140354.  
PR 24-JUN-1999: 99US-0140355.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142056.  
PR 06-JUL-1999: 99US-0142390.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142920.  
PR 12-JUL-1999: 99US-0143777.  
PR 13-JUL-1999: 99US-0143842.  
PR 14-JUL-1999: 99US-0143824.  
PR 15-JUL-1999: 99US-0144006.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144632.

PR 20-JUL-1999: 99US-0144884.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0145086.  
PR 21-JUL-1999: 99US-0145088.  
PR 22-JUL-1999: 99US-0145085.  
PR 23-JUL-1999: 99US-0145087.  
PR 23-JUL-1999: 99US-0145089.  
PR 23-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 24-JUL-1999: 99US-0145224.  
PR 24-JUL-1999: 99US-0145276.  
PR 26-JUL-1999: 99US-0145913.  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 01-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147302.  
PR 04-AUG-1999: 99US-0147192.  
PR 05-AUG-1999: 99US-0147260.  
PR 05-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 06-AUG-1999: 99US-0147493.  
PR 08-AUG-1999: 99US-0147935.  
PR 10-AUG-1999: 99US-0148171.  
PR 11-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148341.  
PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148584.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149175.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149929.  
PR 20-AUG-1999: 99US-0149930.  
PR 23-AUG-1999: 99US-0149930.  
PR 25-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151303.  
PR 31-AUG-1999: 99US-0151338.  
PR 01-SEP-1999: 99US-0151930.  
PR 07-SEP-1999: 99US-0152363.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 15-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 24-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155489.  
PR 28-SEP-1999: 99US-0156458.  
PR 29-SEP-1999: 99US-0156596.  
PR 04-OCT-1999: 99US-0157117.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158432.  
PR 13-OCT-1999: 99US-0158443.  
PR 13-OCT-1999: 99US-0159293.  
PR 13-OCT-1999: 99US-0159294.  
PR 14-OCT-1999: 99US-0159495.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159637.





PP	04-OCT-1999:	990S-0157713.
PR	05-OCT-1999:	990S-0157753.
PR	06-OCT-1999:	990S-0157865.
PR	07-OCT-1999:	990S-0158029.
PR	08-OCT-1999:	992S-0158232.
PR	12-OCT-1999:	990S-0158369.
PR	13-OCT-1999:	990S-0159293.
PR	13-OCT-1999:	990S-0159294.
PR	13-OCT-1999:	990S-0159295.
PR	14-OCT-1999:	990S-0159329.
PR	14-OCT-1999:	990S-0159330.
PR	14-OCT-1999:	990S-0159331.
PR	14-OCT-1999:	990S-0159637.
PR	14-OCT-1999:	990S-0159638.
PR	18-OCT-1999:	990S-0159584.
PR	21-OCT-1999:	990S-0160741.
PR	21-OCT-1999:	990S-0160767.
PR	21-OCT-1999:	990S-0160768.
PR	21-OCT-1999:	990S-0160770.
PR	21-OCT-1999:	990S-0160815.
PR	21-OCT-1999:	990S-0160815.
PR	22-OCT-1999:	990S-0160980.
PR	22-OCT-1999:	990S-0160981.
PR	22-OCT-1999:	990S-0160984.
PR	25-OCT-1999:	990S-0161404.
PR	25-OCT-1999:	990S-0161405.
PR	26-OCT-1999:	990S-0161406.
PR	26-OCT-1999:	990S-0161359.
PR	26-OCT-1999:	990S-0161360.
PR	26-OCT-1999:	990S-0161361.
PR	26-OCT-1999:	990S-0261920.
PR	26-OCT-1999:	990S-0161992.
PR	26-OCT-1999:	990S-0161993.
PR	29-OCT-1999:	990S-0162142.
Query Match: 37 29; Score 687; DB 21; Length 379;		
Post Local Similarity 40 2%; Prev No 1; %=60;		
Matches 144; Conservative 70; Mismatches 128; Indels 16; Gaps 6;		
QY	2 MNQEFSSYAQNSSFPTQVWASMAAPAL-EVAVELLSKPHNLDLAAALDCAAGPNTF 60	
DH	31 msqgdgansysansrqrkvllemakprlvyrnlccmmmlldfrcylkvaeldgsseqnsf 89	
QY	61 AVISTIRPMMEKKRPRIKQTGFICVTINILFGNDPNTIKRQ----SSPVGNPCEEVP 116	
DH	90 lafellntlnvlgcvkvkspeidccclndrpndfnfkfyrfinkelmtnkss--- 146	
QY	117 CYVMWVWSHSLSHLEFNSSLHVSSHVVHHMLDAPKLDSKEFLANLKRLISKSPR 176	
DH	147 cfrvgdgagstysrltfirnszllhszjaldhlskr-----ekleknajdlhsbsfq 200	
QY	177 VWRPAVI SQPHNPEFTIMNAKSQEVVRWDQNVLLIQRNG-SQPSIMQSSPFMTAMT 235	
DH	201 s-q-h-g-j-l-a-p-q-k-l-f-m-o-l-l-e-v-s-g-j-u-t-l-f-i-g-u-d-l-y-d-c-h-l-l-s-e-d 260	
QY	236 AEVSCLLDECKLDFENIPRSFALEVEKFLVERTSFTTEHNESCLDSVMCEMKR 295	
DH	261 rdlvtelgvscslsdltmrrfydpmvqelkvqkcsafsefneshtladyyecdaf 320	
QY	296 VPGREKFKVAFATPTINSQGPETIMKLVNKRFHIVNSDLAEAKLFRTSIILVSK 353	
DH	321 g-ap-n-e-n-a-m-j-r-a-v-s-p-m-l-a-b-t-g-e-c-l-l-d-l-l-k-y-a-v-t-r-a-n-c-r-i-k-l-v-s-l-v-s-l-k 378	
RESULT 8		
ID	AAG37439 standard; Protein: 379 AA.	
AC	AAG37439;	
DT	18-OCT-2000 (first entry)	
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 46035.	

XX protein identification, signal transduction pathway; metabolism pathway;  
KM hybridization assay, genetic mapping, gene expression control; promoter;  
KM termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000: 2000EP-0301439  
XX  
XX 25-FEB-1999: 99US-0121825.  
XX 05-MAR-1999: 99US-0123180.  
XX 09-MAY-1999: 99US-0123548.  
XX 23-MAR-1999: 99US-0125788.  
XX 25-MAR-1999: 99US-0126264.  
XX 29-MAR-1999: 99US-0126785.  
XX 01-APR-1999: 99US-0127452.  
XX 06-APR-1999: 99US-0128234.  
XX 08-APR-1999: 99US-0128714.  
XX 16-APR-1999: 99US-0129845.  
XX 19-APR-1999: 99US-0130077.  
XX 21-APR-1999: 99US-0130449.  
XX 23-APR-1999: 99US-0130510.  
XX 28-APR-1999: 99US-0131449.  
XX 30-APR-1999: 99US-0132048.  
XX 30-APR-1999: 99US-0132407.  
XX 04-MAY-1999: 99US-0132484.  
XX 05-MAY-1999: 99US-0132485.  
XX 06-MAY-1999: 99US-0132486.  
XX 07-MAY-1999: 99US-0132487.  
XX 11-MAY-1999: 99US-0132863.  
XX 14-MAY-1999: 99US-0134256.  
XX 14-MAY-1999: 99US-0134218.  
XX 14-MAY-1999: 99US-0134219.  
XX 14-MAY-1999: 99US-0134221.  
XX 18-MAY-1999: 99US-0134370.  
XX 19-MAY-1999: 99US-0134768.  
XX 20-MAY-1999: 99US-0134941.  
XX 21-MAY-1999: 99US-0135124.  
XX 21-MAY-1999: 99US-0135353.  
XX 24-MAY-1999: 99US-0135629.  
XX 25-MAY-1999: 99US-0136021.  
XX 27-MAY-1999: 99US-0136392.  
XX 28-MAY-1999: 99US-0136782.  
XX 01-JUN-1999: 99US-0137222.  
XX 03-JUN-1999: 99US-0137528.  
XX 04-JUN-1999: 99US-0137602.  
XX 07-JUN-1999: 99US-0137724.  
XX 08-JUN-1999: 99US-0138094.  
XX 10-JUN-1999: 99US-0138540.  
XX 10-JUN-1999: 99US-0138847.  
XX 14-JUN-1999: 99US-0139119.  
XX 16-JUN-1999: 99US-0139452.  
XX 16-JUN-1999: 99US-0139453.  
XX 17-JUN-1999: 99US-0139492.  
XX 18-JUN-1999: 99US-0139454.  
XX 18-JUN-1999: 99US-0139455.  
XX 18-JUN-1999: 99US-0139456.  
XX 18-JUN-1999: 99US-0139457.  
XX 18-JUN-1999: 99US-0139458.  
XX 18-JUN-1999: 99US-0139459.  
XX 18-JUN-1999: 99US-0139460.  
XX 18-JUN-1999: 99US-0139461.  
XX 18-JUN-1999: 99US-0139462.  
XX 18-JUN-1999: 99US-0139463.  
XX 18-JUN-1999: 99US-0139750.  
XX 18-JUN-1999: 99US-0139763.  
XX 21-JUN-1999: 99US-0139817.  
XX 22-JUN-1999: 99US-0139899.  
XX 23-JUN-1999: 99US-0140353.  
XX 24-JUN-1999: 99US-0140354.  
XX 28-JUN-1999: 99US-0140695.  
XX 28-JUN-1999: 99US-0140823.  
XX 29-JUN-1999: 99US-0140991.  
XX 30-JUN-1999: 99US-0141287.  
XX 01-JUL-1999: 99US-0141842.  
XX 01-JUL-1999: 99US-0142154.  
XX 02-JUL-1999: 99US-0142965.  
XX 06-JUL-1999: 99US-0142969.  
XX 08-JUL-1999: 99US-0142803.  
XX 09-JUL-1999: 99US-0142920.  
XX 12-JUL-1999: 99US-0142977.  
XX 13-JUL-1999: 99US-0143542.  
XX 14-JUL-1999: 99US-0143634.  
XX 15-JUL-1999: 99US-0144005.  
XX 16-JUL-1999: 99US-0144085.  
XX 16-JUL-1999: 99US-0144086.  
XX 19-JUL-1999: 99US-0144325.  
XX 19-JUL-1999: 99US-0144331.  
XX 19-JUL-1999: 99US-0144332.  
XX 19-JUL-1999: 99US-0144333.  
XX 19-JUL-1999: 99US-0144334.  
XX 19-JUL-1999: 99US-0144335.  
XX 20-JUL-1999: 99US-0144352.  
XX 20-JUL-1999: 99US-0144432.  
XX 20-JUL-1999: 99US-0144884.  
XX 21-JUL-1999: 99US-0144814.  
XX 21-JUL-1999: 99US-0145086.  
XX 21-JUL-1999: 99US-0145088.  
XX 22-JUL-1999: 99US-0145087.  
XX 22-JUL-1999: 99US-0145087.  
XX 22-JUL-1999: 99US-0145089.  
XX 23-JUL-1999: 99US-0145192.  
XX 23-JUL-1999: 99US-0145145.  
XX 23-JUL-1999: 99US-0145218.  
XX 23-JUL-1999: 99US-0145224.  
XX 26-JUL-1999: 99US-0145376.  
XX 27-JUL-1999: 99US-0145913.  
XX 27-JUL-1999: 99US-0145918.  
XX 27-JUL-1999: 99US-0145919.  
XX 28-JUL-1999: 99US-0145951.  
XX 02-AUG-1999: 99US-0146386.  
XX 02-AUG-1999: 99US-0146388.  
XX 02-AUG-1999: 99US-0146389.  
XX 03-AUG-1999: 99US-0147038.  
XX 04-AUG-1999: 99US-0147294.  
XX 04-AUG-1999: 99US-0147302.  
XX 05-AUG-1999: 99US-0147392.  
XX 06-AUG-1999: 99US-0147260.  
XX 06-AUG-1999: 99US-0147303.  
XX 06-AUG-1999: 99US-0147416.  
XX 06-AUG-1999: 99US-0147493.  
XX 09-AUG-1999: 99US-0147935.  
XX 10-AUG-1999: 99US-0148171.  
XX 11-AUG-1999: 99US-0148319.  
XX 12-AUG-1999: 99US-0148341.  
XX 13-AUG-1999: 99US-0148565.  
XX 13-AUG-1999: 99US-0148684.  
XX 16-AUG-1999: 99US-0149368.  
XX 17-AUG-1999: 99US-0149175.  
XX 18-AUG-1999: 99US-0149426.  
XX 20-AUG-1999: 99US-0149722.  
XX 20-AUG-1999: 99US-0149723.  
XX 20-AUG-1999: 99US-0149929.  
XX 23-AUG-1999: 99US-0149902.  
XX 23-AUG-1999: 99US-0149929.  
XX 25-AUG-1999: 99US-0150566.  
XX 26-AUG-1999: 99US-0150884.  
XX 27-AUG-1999: 99US-0151065.  
XX 27-AUG-1999: 99US-0151066.  
XX 27-AUG-1999: 99US-0151080.  
XX 30-AUG-1999: 99US-0151303.

```

PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151430.
PR 07-SEP-1999; 9905-0152363.
PR 10-SEP-1999; 9905-0153070.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154018.
PR 16-SEP-1999; 9905-0154036.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155559.
PR 28-SEP-1999; 9905-0156458.
PR 29-SEP-1999; 9905-0156596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158232.
PR 12-OCT-1999; 9905-0158369.
PR 13-OCT-1999; 9905-0159293.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159637.
PR 14-OCT-1999; 9905-0159638.
PR 18-OCT-1999; 9905-0159584.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160778.
PR 21-OCT-1999; 9905-0160779.
PR 21-OCT-1999; 9905-0160814.
PR 21-OCT-1999; 9905-0160815.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161450.
PR 26-OCT-1999; 9905-0161361.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161920.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

Query Map: 37-38: Score 687: 106-21: Length 473:  
 Best Local Similarity: 40-28: Prod No: 158-60:  
 Matches: 144: Conservative: 70: Mismatches: 128: Indels: 16: Gaps: 6:

```

CY 2 MNRKQSSSTAAVNSSTLVVAVSMAQVAF--LNAYVILSKDFHQAIAADLQAGANLIF 60
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 31 msqdyasysaustlkkkylsmakvlyvnteeamndiplylkkvelqssysustf 89
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 61 AVISTIRMMKRRKPRKNCGLFLVYVNLNLPNDNFILKEL----SSVGVNKCEEVF 119
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 90 lafeifnlivtoehvknkspedvclodlprodfatctfkvffnkeimimnss--- 146
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 117 GYVAVPSSSHLPENSHI VHSYSVHMLTVAKKLSPEPLAI NCKIVISKSHF 179
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 147 cfvygagpsfysrlfssstlhlssyaltklskyl--- oklmmknkgnlyrfsespa 200
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 177 VVRKAVLSQHEFTMLLNKSSQEVVNDIMVILKQVQ--SNQSIWQSTFWELLAMAI 245
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 201 saykaylnqtkdftmlrlrscvlsngmrvlftfgmrladpdyrccchwtllnsal 260
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 236 AELVSQGLIDDKLDITNINISFASLEEKDIVERDGSFTIDIEEGPDDSDVMQNDKW 295
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
DB 261 rdlrtfeglvyskkladfrimfydwnvclkerlqkcdsflmcdesgldlqlyggedf 320
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |
CY 296 VRGKRFKVVAPTEPILSNQPFPEIMDKLYRFTIILVSDLEAKLPKTSILLVISK 353
    . . . . . | . . . . . | . . . . . | . . . . . | . . . . . |

```

```

DB 321 cagmcaanglravsepmllshqeeidldldkyayhvgkanorokitvslvstlk 378
RESULT 9
ID AAG12688 standard; Protein: 327 AA.
XX AAG12688.
XX AC AAG12688.
XX DE 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 11899.
XX DE Protein identification: signal transduction pathway; metabolic pathway;
XX DE hybridisation assay; genetic mapping; gene expression control; precursor;
XX DE termination sequence.
XX US Arabidopsis thaliana.
XX PN EP1031405-A2.
XX PD 06-SEP-2000.
XX PE 25-FEB-2000; 2000EP-00014-9
XX XX 25-FEB-1999; 9905-0121825.
XX XX 05-MAY-1999; 9905-0123180.
XX XX 04-MAR-1999; 9905-0123548.
XX XX 23-MAR-1999; 9905-0125788.
XX XX 25-MAR-1999; 9905-0126264.
XX XX 29-MAR-1999; 9905-0126785.
XX XX 01-APR-1999; 9905-0127462.
XX XX 06-APR-1999; 9905-0128234.
XX XX 08-APR-1999; 9905-0128714.
XX XX 16-APR-1999; 9905-0129845.
XX XX 19-APR-1999; 9905-0130077.
XX XX 21-APR-1999; 9905-0130449.
XX XX 23-APR-1999; 9905-0130510.
XX XX 23-APR-1999; 9905-0130891.
XX XX 28-APR-1999; 9905-0131449.
XX XX 30-APR-1999; 9905-0132048.
XX XX 30-APR-1999; 9905-0132407.
XX XX 04-MAY-1999; 9905-0132484.
XX XX 05-MAY-1999; 9905-0132485.
XX XX 06-MAY-1999; 9905-0132486.
XX XX 06-MAY-1999; 9905-0132487.
XX XX 07-MAY-1999; 9905-0132863.
XX XX 11-MAY-1999; 9905-0134256.
XX XX 14-MAY-1999; 9905-0134218.
XX XX 14-MAY-1999; 9905-0134219.
XX XX 14-MAY-1999; 9905-0134221.
XX XX 14-MAY-1999; 9905-0134370.
XX XX 16-MAY-1999; 9905-0134768.
XX XX 19-MAY-1999; 9905-0135124.
XX XX 20-MAY-1999; 9905-0135124.
XX XX 24-MAY-1999; 9905-0135259.
XX XX 25-MAY-1999; 9905-0135021.
XX XX 27-MAY-1999; 9905-0136392.
XX XX 28-MAY-1999; 9905-0136782.
XX XX 01-JUN-1999; 9905-0137222.
XX XX 04-JUN-1999; 9905-0137228.
XX XX 04-JUN-1999; 9905-0137522.
XX XX 07-JUN-1999; 9905-0137724.
XX XX 08-JUN-1999; 9905-0138094.
XX XX 10-JUN-1999; 9905-0138450.
XX XX 10-JUN-1999; 9905-0138847.
XX XX 14-JUN-1999; 9905-0139119.
XX XX 16-JUN-1999; 9905-0139452.
XX XX 16-JUN-1999; 9905-0139453.
XX XX 17-JUN-1999; 9905-0139492.
XX XX 18-JUN-1999; 9905-0139454.

```

```

PR 18-JUN-1999: 990S-0139455-
PR 18-JUN-1999: 990S-0139456-
PR 18-JUN-1999: 990S-0139457-
PR 18-JUN-1999: 990S-0139458-
PR 18-JUN-1999: 990S-0139459-
PR 18-JUN-1999: 990S-0139460-
PR 18-JUN-1999: 990S-0139461-
PR 18-JUN-1999: 990S-0139462-
PR 18-JUN-1999: 990S-0139463-
PR 18-JUN-1999: 990S-0139750-
PR 18-JUN-1999: 990S-0139763-
PR 21-JUN-1999: 990S-0139817-
PR 22-JUN-1999: 990S-0139899-
PR 21-JUN-1999: 990S-0140353-
PR 21-JUN-1999: 990S-0140354-
PR 24-JUN-1999: 990S-0140605-
PR 28-JUN-1999: 990S-0140823-
PR 29-JUN-1999: 990S-0140941-
PR 30-JUN-1999: 990S-0141287-
PR 01-JUL-1999: 990S-0141842-
PR 01-JUL-1999: 990S-0142154-
PR 02-JUL-1999: 990S-0142055-
PR 06-JUL-1999: 990S-0142390-
PR 08-JUL-1999: 990S-0142803-
PR 09-JUL-1999: 990S-0142920-
PR 12-JUL-1999: 990S-0142977-
PR 13-JUL-1999: 990S-0143542-
PR 14-JUL-1999: 990S-0143624-
PR 15-JUL-1999: 990S-0144005-
PR 16-JUL-1999: 990S-0144085-
PR 16-JUL-1999: 990S-0144086-
PR 19-JUL-1999: 990S-0144325-
PR 19-JUL-1999: 990S-0144331-
PR 19-JUL-1999: 990S-0144332-
PR 19-JUL-1999: 990S-0144333-
PR 19-JUL-1999: 990S-0144334-
PR 19-JUL-1999: 990S-0144335-
PR 20-JUL-1999: 990S-0144352-
PR 20-JUL-1999: 990S-0144632-
PR 20-JUL-1999: 990S-0144884-
PR 21-JUL-1999: 990S-0144814-
PR 21-JUL-1999: 990S-0145086-
PR 21-JUL-1999: 990S-0145087-
PR 22-JUL-1999: 990S-0145085-
PR 22-JUL-1999: 990S-0145087-
PR 22-JUL-1999: 990S-0145089-
PR 22-JUL-1999: 990S-0145192-
PR 23-JUL-1999: 990S-0145197-
PR 23-JUL-1999: 990S-0145218-
PR 23-JUL-1999: 990S-0145224-
PR 26-JUL-1999: 990S-0145276-
PR 27-JUL-1999: 990S-0145913-
PR 27-JUL-1999: 990S-0145918-
PR 27-JUL-1999: 990S-0145919-
PR 28-JUL-1999: 990S-0145951-
PR 02-AUG-1999: 990S-0146386-
PR 02-AUG-1999: 990S-0146388-
PR 02-AUG-1999: 990S-0146389-
PR 03-AUG-1999: 990S-0147038-
PR 04-AUG-1999: 990S-0147204-
PR 04-AUG-1999: 990S-0147302-
PR 05-AUG-1999: 990S-0147192-
PR 05-AUG-1999: 990S-0147260-
PR 06-AUG-1999: 990S-0147303-
PR 06-AUG-1999: 990S-0147416-
PR 09-AUG-1999: 990S-0147933-
PR 09-AUG-1999: 990S-0147935-
PR 10-AUG-1999: 990S-0148171-
PR 10-AUG-1999: 990S-0148319-
PR 12-AUG-1999: 990S-0148341-
PR 13-AUG-1999: 990S-0148565-
PR 13-AUG-1999: 990S-0148684-
PR 16-AUG-1999: 990S-0149368-

```

```

PR 17-AUG-1999: 990S-0149375-
PR 18-AUG-1999: 990S-0149426-
PR 20-AUG-1999: 990S-0149722-
PR 20-AUG-1999: 990S-0149723-
PR 20-AUG-1999: 990S-0149929-
PR 23-AUG-1999: 990S-0149902-
PR 23-AUG-1999: 990S-0149930-
PR 25-AUG-1999: 990S-0150666-
PR 25-AUG-1999: 990S-0150884-
PR 27-AUG-1999: 990S-0151065-
PR 27-AUG-1999: 990S-0151066-
PR 27-AUG-1999: 990S-0151080-
PR 30-AUG-1999: 990S-0151303-
PR 31-AUG-1999: 990S-0151438-
PR 01-SEP-1999: 990S-0151930-
PR 07-SEP-1999: 990S-0152363-
PR 10-SEP-1999: 990S-0154070-
PR 13-SEP-1999: 990S-0153758-
PR 15-SEP-1999: 990S-0154018-
PR 16-SEP-1999: 990S-0154039-
PR 20-SEP-1999: 990S-0154779-
PR 22-SEP-1999: 990S-0155133-
PR 23-SEP-1999: 990S-0155486-
PR 24-SEP-1999: 990S-0155659-
PR 28-SEP-1999: 990S-0156458-
PR 29-SEP-1999: 990S-0156596-
PR 04-OCT-1999: 990S-0157117-
PR 05-OCT-1999: 990S-0157753-
PR 06-OCT-1999: 990S-0157865-
PR 07-OCT-1999: 990S-0158029-
PR 09-OCT-1999: 990S-0158232-
PR 12-OCT-1999: 990S-0158369-
PR 13-OCT-1999: 990S-0159293-
PR 13-OCT-1999: 990S-0159294-
PR 13-OCT-1999: 990S-0159295-
PR 14-OCT-1999: 990S-0159329-
PR 14-OCT-1999: 990S-0159330-
PR 14-OCT-1999: 990S-0159331-
PR 14-OCT-1999: 990S-0159637-
PR 14-OCT-1999: 990S-0159638-
PR 18-OCT-1999: 990S-0159584-
PR 21-OCT-1999: 990S-0160741-
PR 21-OCT-1999: 990S-0160767-
PR 21-OCT-1999: 990S-0160768-
PR 21-OCT-1999: 990S-0160770-
PR 21-OCT-1999: 990S-0160814-
PR 21-OCT-1999: 990S-0160815-
PR 21-OCT-1999: 990S-0160948-
PR 22-OCT-1999: 990S-0160981-
PR 22-OCT-1999: 990S-0160989-
PR 25-OCT-1999: 990S-0161404-
PR 25-OCT-1999: 990S-0161405-
PR 25-OCT-1999: 990S-0161406-
PR 26-OCT-1999: 990S-0161359-
PR 26-OCT-1999: 990S-0161360-
PR 26-OCT-1999: 990S-0161361-
PR 28-OCT-1999: 990S-0161923-
PR 28-OCT-1999: 990S-0161922-
PR 28-OCT-1999: 990S-0161993-
PR 29-OCT-1999: 990S-0162142-

```

Query Match: 34.7%, Score 641, LB 21, Length 327;  
 Best Local Similarity: 40.2%, Pred. No. 4,96-56;  
 Matches 135; Conservative 64; Mismatches 121; Indels 16; Gaps 6;

```

QY 24 MAFPAL-ENAVETLPSFPHICALNAADLCZACHTREFAVETIKRMKREKCELNQT, 82
DB 1 makp:lvrtccommmlidpvt-krpccqcsqststaltatellimlvrttqpmckasp 59
QY 83 ETGVNINLFCNIFNLKXZL SSIVYGNSTV-VQYVWGVH-ZSF-RCHLPPNPSTHL, 138
DB 60 eldcclndipndfntllkfvlllnkclntlkksr cltygqapstyslstrshl 116

```





PR 07-MAY-1999: 990S-0132863  
PR 11-MAY-1999: 990S-0134256  
PR 14-MAY-1999: 990S-0134218  
PR 14-MAY-1999: 990S-0134219  
PR 14-MAY-1999: 990S-0134221  
PR 14-MAY-1999: 990S-0134370  
PR 18-MAY-1999: 990S-0134768  
PR 19-MAY-1999: 990S-0134941  
PR 20-MAY-1999: 990S-0135124  
PR 21-MAY-1999: 990S-0135353  
PR 24-MAY-1999: 990S-0135529  
PR 25-MAY-1999: 990S-0136021  
PR 27-MAY-1999: 990S-0136392  
PR 28-MAY-1999: 990S-0136732  
PR 01-JUN-1999: 990S-0137222  
PR 03-JUN-1999: 990S-0137528  
PR 04-JUN-1999: 990S-0137502  
PR 07-JUN-1999: 990S-0137724  
PR 08-JUN-1999: 990S-0138094  
PR 10-JUN-1999: 990S-0138540  
PR 10-JUN-1999: 990S-0138847  
PR 14-JUN-1999: 990S-0139119  
PR 16-JUN-1999: 990S-0139452  
PR 16-JUN-1999: 990S-0139453  
PR 17-JUN-1999: 990S-0139492  
PR 18-JUN-1999: 990S-0139454  
PR 18-JUN-1999: 990S-0139455  
PR 18-JUN-1999: 990S-0139456  
PR 18-JUN-1999: 990S-0139457  
PR 18-JUN-1999: 990S-0139458  
PR 18-JUN-1999: 990S-0139459  
PR 18-JUN-1999: 990S-0139460  
PR 18-JUN-1999: 990S-0139461  
PR 18-JUN-1999: 990S-0139462  
PR 18-JUN-1999: 990S-0139463  
PR 18-JUN-1999: 990S-0139750  
PR 18-JUN-1999: 990S-0139751  
PR 21-JUN-1999: 990S-0139817  
PR 22-JUN-1999: 990S-0139899  
PR 23-JUN-1999: 990S-0140354  
PR 23-JUN-1999: 990S-0140354  
PR 24-JUN-1999: 990S-0140655  
PR 26-JUN-1999: 990S-0140693  
PR 29-JUN-1999: 990S-0140991  
PR 30-JUN-1999: 990S-0141287  
PR 01-JUL-1999: 990S-0141842  
PR 01-JUL-1999: 990S-0142154  
PR 02-JUL-1999: 990S-0142356  
PR 06-JUL-1999: 990S-0142390  
PR 06-JUL-1999: 990S-0142403  
PR 09-JUL-1999: 990S-0142440  
PR 12-JUL-1999: 990S-0142977  
PR 13-JUL-1999: 990S-0143542  
PR 14-JUL-1999: 990S-0143624  
PR 15-JUL-1999: 990S-0144005  
PR 16-JUL-1999: 990S-0144085  
PR 16-JUL-1999: 990S-0144086  
PR 19-JUL-1999: 990S-0144325  
PR 19-JUL-1999: 990S-0144331  
PR 19-JUL-1999: 990S-0144332  
PR 19-JUL-1999: 990S-0144333  
PR 19-JUL-1999: 990S-0144334  
PR 19-JUL-1999: 990S-0144335  
PR 20-JUL-1999: 990S-0144352  
PR 20-JUL-1999: 990S-0144352  
PR 20-JUL-1999: 990S-0144352  
PR 21-JUL-1999: 990S-0144414  
PR 21-JUL-1999: 990S-0145086  
PR 21-JUL-1999: 990S-0145088  
PR 22-JUL-1999: 990S-0145085  
PR 22-JUL-1999: 990S-0145087  
PR 22-JUL-1999: 990S-0145089  
PR 22-JUL-1999: 990S-0145192

PR 23-JUL-1999: 990S-0145145  
PR 23-JUL-1999: 990S-0145218  
PR 23-JUL-1999: 990S-0145224  
PR 26-JUL-1999: 990S-0145276  
PR 27-JUL-1999: 990S-0145913  
PR 27-JUL-1999: 990S-0145918  
PR 27-JUL-1999: 990S-0145919  
PR 28-JUL-1999: 990S-0145951  
PR 02-AUG-1999: 990S-0146388  
PR 02-AUG-1999: 990S-0146389  
PR 03-AUG-1999: 990S-0147038  
PR 04-AUG-1999: 990S-0147204  
PR 04-AUG-1999: 990S-0147204  
PR 05-AUG-1999: 990S-0147192  
PR 05-AUG-1999: 990S-0147260  
PR 06-AUG-1999: 990S-0147303  
PR 06-AUG-1999: 990S-0147416  
PR 09-AUG-1999: 990S-0147493  
PR 09-AUG-1999: 990S-0147935  
PR 10-AUG-1999: 990S-0148171  
PR 11-AUG-1999: 990S-0148319  
PR 12-AUG-1999: 990S-0148341  
PR 13-AUG-1999: 990S-0148565  
PR 13-AUG-1999: 990S-0148584  
PR 16-AUG-1999: 990S-0149368  
PR 17-AUG-1999: 990S-0149175  
PR 18-AUG-1999: 990S-0149426  
PR 20-AUG-1999: 990S-0149722  
PR 20-AUG-1999: 990S-0149723  
PR 20-AUG-1999: 990S-0149929  
PR 23-AUG-1999: 990S-0149902  
PR 24-AUG-1999: 990S-0149930  
PR 25-AUG-1999: 990S-0150566  
PR 26-AUG-1999: 990S-0150884  
PR 27-AUG-1999: 990S-0151065  
PR 27-AUG-1999: 990S-0151066  
PR 27-AUG-1999: 990S-0151080  
PR 28-AUG-1999: 990S-0151403  
PR 28-AUG-1999: 990S-0151438  
PR 01-SEP-1999: 990S-0151930  
PR 07-SEP-1999: 990S-0152363  
PR 10-SEP-1999: 990S-0153070  
PR 13-SEP-1999: 990S-0153758  
PR 15-SEP-1999: 990S-0154018  
PR 16-SEP-1999: 990S-0154039  
PR 20-SEP-1999: 990S-0154779  
PR 22-SEP-1999: 990S-0155149  
PR 23-SEP-1999: 990S-0155486  
PR 24-SEP-1999: 990S-0155659  
PR 28-SEP-1999: 990S-0156458  
PR 29-SEP-1999: 990S-0156596  
PR 04-OCT-1999: 990S-0157117  
PR 05-OCT-1999: 990S-0157753  
PR 06-OCT-1999: 990S-0157865  
PR 07-OCT-1999: 990S-0158029  
PR 08-OCT-1999: 990S-0158232  
PR 12-OCT-1999: 990S-0158369  
PR 13-OCT-1999: 990S-0159293  
PR 13-OCT-1999: 990S-0159294  
PR 13-OCT-1999: 990S-0159295  
PR 14-OCT-1999: 990S-0159429  
PR 14-OCT-1999: 990S-0159430  
PR 14-OCT-1999: 990S-0159431  
PR 14-OCT-1999: 990S-0159637  
PR 14-OCT-1999: 990S-0159638  
PR 18-OCT-1999: 990S-0159584  
PR 21-OCT-1999: 990S-0160741  
PR 21-OCT-1999: 990S-0160767  
PR 21-OCT-1999: 990S-0160768  
PR 21-OCT-1999: 990S-0160770  
PR 21-OCT-1999: 990S-0160814  
PR 21-OCT-1999: 990S-0160815

```

PR 22 OCT-1999; 9905-0160989;
PR 22-OCT-1999; 9905-0160981;
PR 22-OCT-1999; 9905-0160989;
PR 25-OCT-1999; 9905-0161404;
PR 25-OCT-1999; 9905-0161405;
PR 25-OCT-1999; 9905-0161406;
PR 26-OCT-1999; 9905-0161359;
PR 26-OCT-1999; 9905-0161360;
PR 26-OCT-1999; 9905-0161361;
PR 28-OCT-1999; 9905-0161360;
PR 28-OCT-1999; 9905-0161361;
PR 28-OCT-1999; 9905-0161362;
PR 28-OCT-1999; 9905-0161363;
PR 29-OCT-1999; 9905-0162142;

Query Match: 26.08, Score 481, Db 21; Length 374;
Best Local Similarity 32.08, P-val 7.4e-40;
Matches 125, Conserved 56, Mismatches 134, Indels 66, Gaps 12;

QY 2 MNKGGESSYAKQNSFTQGVASMAQPALENVETLFSRDFHQA-----LNAAIDGCA 55
DB 9 mkykqkqsyauus ..qagambarsmlhleeel ..evnlhssaspprfcavdlqcs 63
QY 56 GPNFAVSLTKRMKKKELNQCILEQVYLNDLFGNDFTLFGKLSSEVIGKCEEV 115
DB 64 gantvhlidivkhiskrfiaagldpetafatsdpndlnelfq..llpplvsnlcmee- 121
QY 116 PC-----YVMGVGSPFHGFLFPPNSLALVHSSYSVHMLUAVKGLFSRGLALNK 166
DB 122 eciaadqustfvaqyvsfyrilffarldfhsafshwsfysvsvldfrrsaayng 181
QY 167 KIVL---SKTSPVVRVAVLSQFHEDETMELARSGEVVPCNVCLIRGPGSLPSTLM 222
DB 182 rvlhgaqekt-----taykrqgadlaeflraaevkrqgamflvcglgrtsvqptq 236
QY 223 GSC-----FWELLMAIAIMELVSGIDIDEMKILFTIRIPYSLSLEEVKIVPRFS 273
DB 237 ggaqallfghfhdawd-----dlvreglvdaekrdylnlpyaqslqdfkevvdang 289
QY 274 FTIDHIEFPIDLS--VEMGENKMKVGESEKTKVVPFAFTEPTISNGCPLEMLKLYDK-- 328
DB 290 faisklrvykgqgplvnefddasvvgldasstrsvaylvvahlgceelskhlfsvcs 349
QY 329 -----FTHIVVS 335
DB 350 ratshakdvlnqlthivas 370

RESULT 12
AAC6652
ID AAC6652 standard; Protein; 386 AA.
XX
AC AAC6652;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58715
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN PP1031405-A2.
XX
PD 06-SEP-2000
XX
PE 25-FEB-2000, 2000PP-0301439.
XX
PF 25-FEB-1999; 9905-0121825;
PR 05-MAR-1999; 9905-0123180;
PR 09-MAR-1999; 9905-0123548;
PR 23-MAR-1999; 9905-0125788;

```

```

PR 25-MAR-1999; 9905-0125788;
PR 25-MAR-1999; 9905-0126785;
PR 01-APR-1999; 9905-0127462;
PR 06-APR-1999; 9905-0128234;
PR 08-APR-1999; 9905-0128714;
PR 16-APR-1999; 9905-0129845;
PR 19-APR-1999; 9905-0130077;
PR 21-APR-1999; 9905-0130049;
PR 23-APR-1999; 9905-0130510;
PR 23-APR-1999; 9905-0130891;
PR 28-APR-1999; 9905-0131449;
PR 30-APR-1999; 9905-0132048;
PR 30-APR-1999; 9905-0132047;
PR 04-MAY-1999; 9905-0132484;
PR 05-MAY-1999; 9905-0132485;
PR 06-MAY-1999; 9905-0132486;
PR 06-MAY-1999; 9905-0132487;
PR 07-MAY-1999; 9905-0132863;
PR 11-MAY-1999; 9905-0134256;
PR 14-MAY-1999; 9905-0134218;
PR 14-MAY-1999; 9905-0134219;
PR 14-MAY-1999; 9905-0134221;
PR 14-MAY-1999; 9905-0134370;
PR 18-MAY-1999; 9905-0134768;
PR 19-MAY-1999; 9905-0134941;
PR 20-MAY-1999; 9905-0135124;
PR 21-MAY-1999; 9905-0135353;
PR 24-MAY-1999; 9905-0135453;
PR 25-MAY-1999; 9905-0136021;
PR 27-MAY-1999; 9905-0136392;
PR 28-MAY-1999; 9905-0136782;
PR 01-JUN-1999; 9905-0137222;
PR 03-JUN-1999; 9905-0137528;
PR 04-JUN-1999; 9905-0137502;
PR 07-JUN-1999; 9905-0137724;
PR 08-JUN-1999; 9905-0138094;
PR 10-JUN-1999; 9905-0138540;
PR 14-JUN-1999; 9905-0138847;
PR 16-JUN-1999; 9905-0139119;
PR 16-JUN-1999; 9905-0139452;
PR 17-JUN-1999; 9905-0139453;
PR 17-JUN-1999; 9905-0139492;
PR 18-JUN-1999; 9905-0139454;
PR 18-JUN-1999; 9905-0139455;
PR 18-JUN-1999; 9905-0139456;
PR 18-JUN-1999; 9905-0139457;
PR 18-JUN-1999; 9905-0139458;
PR 18-JUN-1999; 9905-0139459;
PR 18-JUN-1999; 9905-0139460;
PR 18-JUN-1999; 9905-0139461;
PR 18-JUN-1999; 9905-0139462;
PR 18-JUN-1999; 9905-0139463;
PR 18-JUN-1999; 9905-0139750;
PR 18-JUN-1999; 9905-0139763;
PR 21-JUN-1999; 9905-0139817;
PR 22-JUN-1999; 9905-0139819;
PR 23-JUN-1999; 9905-0140353;
PR 23-JUN-1999; 9905-0140354;
PR 24-JUN-1999; 9905-0140595;
PR 28-JUN-1999; 9905-0140823;
PR 29-JUN-1999; 9905-0140991;
PR 30-JUN-1999; 9905-0141287;
PR 01-JUL-1999; 9905-0141842;
PR 01-JUL-1999; 9905-0142154;
PR 02-JUL-1999; 9905-0142055;
PR 06-JUL-1999; 9905-0142390;
PR 08-JUL-1999; 9905-0142403;
PR 09-JUL-1999; 9905-0142423;
PR 12-JUL-1999; 9905-0142977;
PR 13-JUL-1999; 9905-0143542;
PR 14-JUL-1999; 9905-0143624;
PR 15-JUL-1999; 9905-0144005;
PR 16-JUL-1999; 9905-0144085;

```



PR 16-JUL-1999: 99US-0144086  
 PR 19-JUL-1999: 99US-0144325  
 PR 19-JUL-1999: 99US-0144331  
 PR 19-JUL-1999: 99US-0144332  
 PR 19-JUL-1999: 99US-0144333  
 PR 19-JUL-1999: 99US-0144334  
 PR 19-JUL-1999: 99US-0144335  
 PR 20-JUL-1999: 99US-0144352  
 PR 20-JUL-1999: 99US-0144352  
 PR 20-JUL-1999: 99US-0144884  
 PR 21-JUL-1999: 99US-0144814  
 PR 21-JUL-1999: 99US-0145086  
 PR 21-JUL-1999: 99US-0145088  
 PR 22-JUL-1999: 99US-0145085  
 PR 22-JUL-1999: 99US-0145097  
 PR 22-JUL-1999: 99US-0145099  
 PR 22-JUL-1999: 99US-0145152  
 PR 23-JUL-1999: 99US-0145145  
 PR 23-JUL-1999: 99US-0145151  
 PR 23-JUL-1999: 99US-0145224  
 PR 26-JUL-1999: 99US-0145276  
 PR 27-JUL-1999: 99US-0145313  
 PR 27-JUL-1999: 99US-0145318  
 PR 27-JUL-1999: 99US-0145319  
 PR 28-JUL-1999: 99US-0145351  
 PR 02-AUG-1999: 99US-0145386  
 PR 02-AUG-1999: 99US-0145388  
 PR 02-AUG-1999: 99US-0145389  
 PR 03-AUG-1999: 99US-0147038  
 PR 04-AUG-1999: 99US-0147204  
 PR 04-AUG-1999: 99US-0147302  
 PR 05-AUG-1999: 99US-0147302  
 PR 05-AUG-1999: 99US-0147360  
 PR 06-AUG-1999: 99US-0147303  
 PR 06-AUG-1999: 99US-0147416  
 PR 06-AUG-1999: 99US-0147463  
 PR 09-AUG-1999: 99US-0147335  
 PR 10-AUG-1999: 99US-0148171  
 PR 11-AUG-1999: 99US-0148319  
 PR 12-AUG-1999: 99US-0148341  
 PR 13-AUG-1999: 99US-0148365  
 PR 13-AUG-1999: 99US-0148365  
 PR 16-AUG-1999: 99US-0149368  
 PR 17-AUG-1999: 99US-0149175  
 PR 18-AUG-1999: 99US-0149426  
 PR 20-AUG-1999: 99US-0149722  
 PR 20-AUG-1999: 99US-0149723  
 PR 20-AUG-1999: 99US-0149729  
 PR 23-AUG-1999: 99US-0149902  
 PR 23-AUG-1999: 99US-0149902  
 PR 25-AUG-1999: 99US-0150366  
 PR 25-AUG-1999: 99US-0150384  
 PR 27-AUG-1999: 99US-0151065  
 PR 27-AUG-1999: 99US-0151065  
 PR 27-AUG-1999: 99US-0151080  
 PR 30-AUG-1999: 99US-0151080  
 PR 31-AUG-1999: 99US-0151138  
 PR 01-SEP-1999: 99US-0151138  
 PR 07-SEP-1999: 99US-0152363  
 PR 10-SEP-1999: 99US-0153070  
 PR 13-SEP-1999: 99US-0153758  
 PR 15-SEP-1999: 99US-0154018  
 PR 16-SEP-1999: 99US-0154039  
 PR 20-SEP-1999: 99US-0154779  
 PR 22-SEP-1999: 99US-0155139  
 PR 23-SEP-1999: 99US-0155486  
 PR 24-SEP-1999: 99US-0155486  
 PR 28-SEP-1999: 99US-0156458  
 PR 29-SEP-1999: 99US-0156596  
 PR 04-OCT-1999: 99US-0157117  
 PR 05-OCT-1999: 99US-0157753  
 PR 06-OCT-1999: 99US-0157865  
 PR 07-OCT-1999: 99US-0158029

PR 08-OCT-1999: 99US-0158232  
 PR 12-OCT-1999: 99US-0158369  
 PR 13-OCT-1999: 99US-0159293  
 PR 13-OCT-1999: 99US-0159294  
 PR 13-OCT-1999: 99US-0159295  
 PR 14-OCT-1999: 99US-0159329  
 PR 14-OCT-1999: 99US-0159330  
 PR 14-OCT-1999: 99US-0159331  
 PR 14-OCT-1999: 99US-0159337  
 PR 14-OCT-1999: 99US-0159368  
 PR 14-OCT-1999: 99US-0159584  
 PR 21-OCT-1999: 99US-0160741  
 PR 21-OCT-1999: 99US-0160767  
 PR 21-OCT-1999: 99US-0160768  
 PR 21-OCT-1999: 99US-0160770  
 PR 21-OCT-1999: 99US-0160814  
 PR 21-OCT-1999: 99US-0160815  
 PR 21-OCT-1999: 99US-0160840  
 PR 21-OCT-1999: 99US-0160881  
 PR 25-OCT-1999: 99US-0161404  
 PR 25-OCT-1999: 99US-0161405  
 PR 25-OCT-1999: 99US-0161406  
 PR 26-OCT-1999: 99US-0161359  
 PR 26-OCT-1999: 99US-0161360  
 PR 26-OCT-1999: 99US-0161361  
 PR 26-OCT-1999: 99US-0161369  
 PR 28-OCT-1999: 99US-0161992  
 PR 28-OCT-1999: 99US-0161993  
 PR 29-OCT-1999: 99US-0162142

Query Match 26.0%, Score 481, DE 21, Length 386,  
 Best Local Similarity 30.8%, Pseq No 7,86,40;  
 Matches 129; Conservative 56; Mismatches 134; Indels 56; Gaps 12;

QY 2 MNPEDESYVANSSTFGVSMAGPALENAVEITFSRPHQA-----LNAELQGA 55  
 DB 21 mkqgkgqdsyans---qagamharsmlhleeit-eenvhnsaasppptavdgss 75  
 QY 56 GPNTPAVTSTIKPKMKKRCPELNCOTPIQVYINLPNNPNTFKKISSEVYGNKCEV 115  
 DB 76 gantvhlidivkhliskrlidaagldpeltatidsldpssldntclq-llpplvslwne- 133  
 QY 116 PC-----YVMGVPGSFHCPLFPNLSLHVSSYVHMLIAPKQISHEALNKG 166  
 DB 134 ecdaadansyvaagysstgtrllparlidlhasfslhwtsqpsvtdtsaaynrg 193  
 QY 167 ELYT---SKSTFVNEAVTSGFEDIMLNPSQEVVNNQVILPPEKSTGSLW 222  
 DB 194 lvtlhgqeklt-----aykltlqdiadellatdaewksgjgallvqgltstvgldq 248  
 QY 224 GSN-----FTWELMALAIKIVSGLIIEKILITINISYASILEKVIIVENGS 273  
 DB 249 gggalllgtthttdawd-----dlvrotgivaokrotgnrnpvapsldtvevdaags 301  
 QY 274 FIDHLEGFELUS--VEMVFNKMWGEKFIKVPATFELIISNGLQPELMKLTAK--- 328  
 DB 302 laidklvvykggsplvynopddasevgrataascsvagvlycoahngcelskflsves 361  
 QY 329 -----FTHLIVVS 335  
 DB 362 ratsbndvynlqfthvas 382

RESULT 13  
 ID AAG46651 standard: Protein; 397 AA.  
 XX AAG46651;  
 AC AAG46651;  
 XX  
 DI 18-OCT-2000 (first entry)  
 XX  
 EE Arabidopsis thaliana protein fragment SEQ ID NO: 58714.











GenForce version 4.5  
 Copyright (c) 1993 - 2000 GenForce, Ltd

GM nucleic acid sequence search, using SW model

Run on: August 31, 2001, 13:27:32 Search time 67.13 seconds  
 (without alignments)  
 4024.246 Million seqs updated/score

Title: US-09-577-657A-2

Perfect score: 1427

Sequence: 1 340 1000 1000 1000 1000

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents NA:  
 1: /seq2\_6/prodata/2/ina/5A.COME seq \*  
 2: /seq2\_6/prodata/2/ina/5A.COME seq \*  
 3: /seq2\_6/prodata/2/ina/5A.COME seq \*  
 4: /seq2\_6/prodata/2/ina/5A.COME seq \*  
 5: /seq2\_6/prodata/2/ina/5A.COME seq \*  
 6: /seq2\_6/prodata/2/ina/5A.COME seq \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52.4	3.7	2421	US-09-027-137-2	Sequence 2, Appl 1
2	52.4	3.7	2421	US-08-630-118A-1	Sequence 1, Appl 1
3	52.4	3.7	2421	US-08-838-399-1	Sequence 1, Appl 1
4	52.4	3.7	2421	US-09-245-819-1	Sequence 1, Appl 1
5	52.4	3.7	2421	US-08-630-118A-1	Sequence 1, Appl 1
6	52.4	3.7	2421	US-08-838-399-1	Sequence 1, Appl 1
7	52.4	3.7	2421	US-09-245-819-1	Sequence 1, Appl 1
8	48.4	3.4	1604	US-09-245-819-1	Sequence 1, Appl 1
9	48.4	3.4	1604	US-08-157-101A-4	Sequence 4, Appl 1
10	46.8	3.3	1393	US-08-174-467-18	Sequence 18, Appl 1
11	46.8	3.3	1393	US-08-452-071-18	Sequence 18, Appl 1
12	46.6	3.3	2147	US-09-014-969-14	Sequence 14, Appl 1
13	46.6	3.3	3581	US-08-738-349-1	Sequence 1, Appl 1
14	46.6	3.3	7218	US-08-232-463-14	Sequence 14, Appl 1
15	46.4	3.3	2852	US-09-063-950-1	Sequence 1, Appl 1
16	46.2	3.2	991	US-08-924-747-25	Sequence 25, Appl 1
17	46.2	3.2	991	US-09-347-338-25	Sequence 25, Appl 1
18	46.2	3.2	991	US-09-347-338-25	Sequence 25, Appl 1
19	46.2	3.2	1600	US-07-861-458C-37	Sequence 37, Appl 1
20	46.2	3.2	3138	US-07-867-196-4	Sequence 4, Appl 1
21	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
22	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
23	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
24	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
25	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
26	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1
27	46.2	3.2	1747	US-08-002-056-4	Sequence 4, Appl 1

28	46	3.2	1737	US-08-801-228-3	Sequence 3, Appl 1
29	46	3.2	1737	US-09-104-296-3	Sequence 3, Appl 1
30	46	3.2	1737	US-08-002-056-4	Sequence 4, Appl 1
31	45.6	3.2	1737	US-08-002-056-4	Sequence 4, Appl 1
32	45.4	3.2	2550	US-08-002-056-4	Sequence 4, Appl 1
33	45.2	3.2	746	US-09-013-810-1	Sequence 1, Appl 1
34	45.2	3.2	746	US-07-867-196-4	Sequence 4, Appl 1
35	45.2	3.2	1552	US-09-413-574-1	Sequence 1, Appl 1
36	45	3.2	1552	US-08-991-426-3	Sequence 3, Appl 1
37	45	3.2	1552	US-08-991-426-3	Sequence 3, Appl 1
38	44.8	3.1	1817	US-08-870-518-8	Sequence 8, Appl 1
39	44.6	3.1	2671	US-08-013-810-1	Sequence 1, Appl 1
40	44	3.1	1602	US-08-630-118A-1	Sequence 1, Appl 1
41	44	3.1	1552	US-08-888-122A-3	Sequence 3, Appl 1
42	44	3.1	1602	US-09-149-879-3	Sequence 3, Appl 1
43	43.6	3.1	334	US-09-032-684-8	Sequence 8, Appl 1
44	43.6	3.1	857	US-08-460-040-1	Sequence 1, Appl 1
45	43.6	3.1	872	US-09-248-335-63	Sequence 63, Appl 1

## ALIGNMENTS

## RESULT 1

US-09-027-137-2

Sequence 2, Application US-09027137

Patent No. 6013450

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Yoo, Henry

TITLE OF INVENTION: CAPI-RELATED PROTEIN

NUMBER OF SEQUENCES: 3

CORRESPONDING ADDRESS:

ADDRESS: 10700 Pharmaceuticals, Inc.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTA for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/027,137

FILING DATE: Filed Herewith

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REFERENCE/DOCKET NUMBER: 36,749

TELEPHONE: 650-855-0555

TELEFAX: 650-845-4166

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 2852 base pairs

TYPE: nucleic acid

STRANDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: PECO0016

CLONE: 2299466

US-09-027-137-2

Query Match: 3.7%

Score 52.6, DB 2

Post Local Similarity: 50.2%

Pred. NO. 0.00015

Length 2852

[illegible]

RESULT 3  
 US-08-838-399-1  
 : Sequence 1, Application US/080838399  
 : Patent No. 5965392  
 :  
 GENERAL INFORMATION:  
 APPLICANT: Hu Ph.D., Yinghe  
 APPLICANT: McCall Ph.D., Michael L.  
 APPLICANT: Bloomquist Ph.D., Brian T.  
 APPLICANT: Flores-Riveros Ph.D., Jaime R.  
 APPLICANT: Cornfield Ph.D., Linda J.  
 TITLE OF INVENTION: Neurexins: X Receptor and Nucleic Acid  
 TITLE OF INVENTION: Sequences  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: McDonnell Hochen Hubbert & Berghoff  
 STREET: 300 South Wacker Drive  
 CITY: Chicago  
 STATE: IL  
 COUNTRY: USA  
 Zip: 60606  
 :  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1 0, Version #1.30  
 :  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08-838-399  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Greenfield Ph.D., Michael S.  
 REGISTRATION NUMBER: 37,147  
 REFERENCE/WORK NUMBER: 96,146/WH 405  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (312)715-1000  
 TELEFAX: (312)715-1234  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 2481 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FEATURE:  
 NAME/KEY: CDS



APPLICANT: Hu Ph D, Yinghe  
APPLICANT: McClell Ph.D., Michael L.  
APPLICANT: Bloomquist Ph.D., Brian T.  
APPLICANT: Flores-Riveros Ph.D., Jaime R.  
APPLICANT: Cornfield Ph.D., Linda J.  
TITLE OF INVENTION: Neuropeptide Y Receptor and Nucleic Acid  
TITLE OF INVENTION: Sequences  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McMahon J Rochem Hulbert & Berghoff  
STREET: 300 South Wacker Drive, 32nd Floor  
CITY: Chicago  
STATE: IL  
COUNTRY: USA  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentLib, Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/40,118A  
FILING DATE: April 8, 1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Greenfield Ph.D., Michael S.  
REGISTRATION NUMBER: 37,142  
REFERENCE/AGENT NUMBER: 95,143/WH 405  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)715-1000









1 NAME: BENT, Stephen A.  
2 REGISTRATION NUMBER: 29, 768  
3 REFERENCE/PROJECT NUMBER: 30472/114 IMM  
4 TELECOMMUNICATION INFORMATION:  
5 TELEPHONE: (703) 836-9400  
6 TELEFAX: (703) 683-4109  
7 TELEX: 899149  
8 INFORMATION FOR SEQ ID NO: 14:  
9 SEQUENCE CHARACTERISTICS:  
10 LENGTH: 7218 base pairs  
11 TYPE: nucleic acid  
12 STRANDEDNESS: single  
13 TOPOLOGY: linear  
14 IMMEDIATE SOURCE:  
15 CLONE: PTZapt-F15  
16 US-08-242-463-14

```

Query Match      3.38; Score 46.6; DB 1; Length 7218;
Host Local Similarity 9.98; Pred. No 0.0082;
Matches 46; Conserved 210; Mismatches 209; Gaps 0.

```

905 gacacatccaatatccacagctatcttgcataccttgaagaatgaataatagtagag 905  
 1406 KRR 1474  
 906 aagacacagatcattccaaatgaatcatatagaagagcttggacttgaagctaaaga 965  
 1466 KRR 1287  
 906 catgacatgataatgagctatgaatggaaagtcttcaacagcttgcagctctccca 1055  
 1385 RRR 1227  
 907 gacacatcattccaaacacagcttggacctggaatcagaatcagaacataatagaacaaatc 1085  
 1226 KRR 1167  
 908 aatccacttctaatttcaatcttgaagaacagacacccaacacacaaatatactcta 1145  
 1166 RRR 1107  
 909 gacatttcacagatctatgataatcttctaagtcttgaagaataaactgattccctatc 1205  
 1106 RRR 1047  
 910 acaatagacacataagagcttcttgcacatgatactgacgaagaag 1250  
 1046 GGAAGCAAGCTCTGCAATTAATCTTTGAGAGCTGATAGGTAAGCAAG 1002

```

1 RESULT 15
2 US-09-0643-950-1
3 : Sequence 1, Application US/09061950C
4 : Patent No. 6225085
5
6 GENERAL INFORMATION:
7
8 APPLICANT: Holzman, Douglas A.
9
10 TITLE OF INVENTION: NINE1 DRSG. PROTEIN AND NUCLEIC ACID MOLECULES AND USES
11
12 TITLE OF INVENTION: THEREFOR
13
14 FILE REFERENCE: MEI-019
15
16 CURRENT APPLICATION NUMBER: US/09/063,950C
17
18 CURRENT FILING DATE: 1998-04-21
19
20 NUMBER OF SEQ ID NOS: 9
21
22 SOFTWARE: PatentIn Ver. 2.0
23
24 SEQ. ID NO. 1
25
26 LENGTH: 2852
27
28 TYPE: DNA
29
30 ORGANISM: Homo sapiens
31
32 FEATURE:

```

```

; NAME/KEY: CDS
; LOCATION: (160)..(2178)
US-09-063-950-1

```

```

Query Match: 3.3% Score 46.4; DB 4; Length 2852;
Host Local Similarity 77.8%; Pred. No. 0.0063;
Matches 56; Conservative 0; Mismatches 16; Indels 0; Gaps 0;
QY 1350 gaatgtgaatgtgaacgtgaattttaaaataaaaataaaataaaataaaat 1409
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2781 gtaagaaaaataaaagatgaagtctgaadaaaaaaaataaaataaaataaa 2840
QY 1410 tcttgacacacac 1421
    | | | | | | |
Db 2841 aagagcgcgcgc 2852

```

Search completed: August 31, 2001, 14:22:18  
Job time: 3286 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM nucleic - nucleic search, using sw model

Run on: August 31, 2001, 13:15:51 : 5037h time 1150.97 seconds (without alignment)

Title: US-09-577-657A-2  
 Perfect score: 1427  
 Sequence: I lgalatcaatgctgttgcag.....atloctggagacggcaattc 1427

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.00

Searched: 10228115 sqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

```
post-processing: Minimum match 0%
                  Maximum match 100%
```

Database :

EST: *	1: qb_est1: *	73: em_estp12: *
2: qb_est2: *	2: qb_est2: *	74: em_estp13: *
3: qb_est3: *	3: qb_est3: *	75: em_estp14: *
4: qb_est4: *	4: qb_est4: *	76: em_estp15: *
5: qb_est5: *	5: qb_est5: *	77: em_estp16: *
6: qb_est6: *	6: qb_est6: *	78: em_estp17: *
7: qb_est7: *	7: qb_est7: *	79: em_estp18: *
8: qb_est8: *	8: qb_est8: *	80: em_estp19: *
9: qb_est9: *	9: qb_est9: *	81: em_estp110: *
10: qb_est10: *	10: qb_est10: *	82: em_estp1: *
11: qb_est11: *	11: qb_est11: *	83: em_estp2: *
12: qb_est12: *	12: qb_est12: *	84: em_estp3: *
13: qb_est13: *	13: qb_est13: *	85: em_estp4: *
14: qb_est14: *	14: qb_est14: *	86: em_estp5: *
15: qb_est15: *	15: qb_est15: *	87: em_estp6: *
16: qb_est16: *	16: qb_est16: *	88: em_estp7: *
17: qb_est17: *	17: qb_est17: *	89: em_estp8: *
18: qb_est18: *	18: qb_est18: *	90: em_estp9: *
19: qb_est19: *	19: qb_est19: *	91: em_estp10: *
20: qb_est20: *	20: qb_est20: *	92: em_estp11: *
21: qb_est21: *	21: qb_est21: *	93: em_estp12: *
22: qb_est22: *	22: qb_est22: *	94: em_estp13: *
23: qb_est23: *	23: qb_est23: *	95: em_estp14: *
24: qb_est24: *	24: qb_est24: *	96: em_estp15: *
25: qb_est25: *	25: qb_est25: *	97: em_estp16: *
26: qb_est26: *	26: qb_est26: *	98: em_estp17: *
27: qb_est27: *	27: qb_est27: *	99: em_estp18: *
28: qb_est28: *	28: qb_est28: *	100: em_estp19: *
29: qb_est29: *	29: qb_est29: *	101: em_estp20: *
30: qb_est30: *	30: qb_est30: *	102: qb_est25: *
31: qb_est31: *	31: qb_est31: *	103: qb_est26: *
32: qb_est32: *	32: qb_est32: *	104: qb_est27: *
33: em_estba: *	33: em_estba: *	105: qb_est28: *
34: em_estba: *	34: em_estba: *	106: qb_est29: *
35: em_estba: *	35: em_estba: *	107: qb_est30: *
36: em_estba: *	36: em_estba: *	108: qb_est31: *
37: em_estba: *	37: em_estba: *	109: qb_est32: *
38: em_estba: *	38: em_estba: *	110: qb_est41: *
39: em_estba: *	39: em_estba: *	111: qb_est42: *
40: em_estba: *	40: em_estba: *	112: qb_est43: *
41: em_estba: *	41: em_estba: *	113: qb_est44: *
42: em_estba: *	42: em_estba: *	114: qb_est45: *
43: em_estba: *	43: em_estba: *	115: qb_est46: *
		116: qb_est47: *

```

117: qb_est14:*
118: qb_est14:*
119: qb_est15:*
120: qb_est15:*
121: qb_est15:*
122: qb_est15:*
123: qb_est15:*
124: qb_est15:*
125: qb_est15:*
126: qb_est15:*
127: qb_est15:*
128: qb_est15:*
129: qb_est15:*
130: qb_est15:*
131: qb_est15:*
132: qb_est15:*
133: qb_est15:*
134: qb_est15:*
135: qb_est15:*
136: qb_est15:*
137: qb_est15:*
138: qb_est15:*
139: qb_est15:*
140: qb_est15:*
141: qb_est15:*
142: qb_est15:*
143: qb_est15:*
144: qb_est15:*
145: qb_est15:*
146: qb_est15:*
147: qb_est15:*
148: qb_est15:*
149: qb_est15:*
150: qb_est15:*
151: qb_est15:*
152: qb_est15:*
153: qb_est15:*
154: qb_est15:*
155: qb_est15:*
156: qb_est15:*
157: qb_est15:*
158: qb_est15:*
159: qb_est15:*
160: qb_est15:*
161: qb_est15:*
162: qb_est15:*
163: qb_est15:*
164: qb_est15:*
165: qb_est15:*
166: qb_est15:*
167: qb_est15:*
168: qb_est15:*
169: qb_est15:*
170: qb_est15:*
171: qb_est15:*
172: qb_est15:*
173: qb_est15:*
174: qb_est15:*
175: qb_est15:*
176: qb_est15:*
177: qb_est15:*
178: qb_est15:*
179: qb_est15:*
180: qb_est15:*
181: qb_est15:*
182: qb_est15:*
183: qb_est15:*
184: qb_est15:*
185: qb_est15:*
186: qb_est15:*
187: qb_est15:*
188: qb_est15:*
189: qb_est15:*

```

---

```

190: qb_est110:*
191: qb_est111:*
192: qb_est111:*
193: qb_est111:*
194: qb_est111:*
195: qb_est111:*
196: qb_est111:*
197: qb_est111:*
198: qb_est111:*
199: qb_est111:*
200: qb_est111:*
201: qb_est111:*
202: qb_est111:*
203: qb_est111:*
204: qb_est111:*
205: qb_est111:*
206: qb_est111:*
207: qb_est111:*
208: qb_est111:*
209: qb_est111:*
210: qb_est111:*
211: qb_est111:*
212: qb_est111:*
213: qb_est111:*
214: qb_est111:*
215: qb_est111:*
216: qb_est111:*
217: qb_est111:*
218: qb_est111:*
219: qb_est111:*
220: qb_est111:*
221: qb_est111:*
222: qb_est111:*
223: qb_est111:*
224: qb_est111:*
225: qb_est111:*
226: qb_est111:*
227: qb_est111:*
228: qb_est111:*
229: qb_est111:*
230: qb_est111:*
231: qb_est111:*
232: qb_est111:*
233: qb_est111:*
234: qb_est111:*
235: qb_est111:*
236: qb_est111:*
237: qb_est111:*
238: qb_est111:*
239: qb_est111:*
240: qb_est111:*
241: qb_est111:*
242: qb_est111:*
243: qb_est111:*
244: qb_est111:*
245: qb_est111:*
246: qb_est111:*
247: qb_est111:*
248: qb_est111:*
249: qb_est111:*
250: qb_est111:*
251: qb_est111:*
252: qb_est111:*
253: qb_est111:*
254: qb_est111:*
255: qb_est111:*
256: qb_est111:*
257: qb_est111:*
258: qb_est111:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.













## AUTHORS

van der Hoeven, R.S., Bezzeredes, J.L., Matero, A.L., Holt, J.E., Jiang  
F., Hansen, T., Craven, M.B., Bowman, C.L., Koning, C.M., Nierman, W.,  
Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue, 4-8 mm buds

## JOURNAL

Unpublished (1999)

## COMMENT

Contact: David Fritsch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4166  
Fax: 864 656 4293  
Email: dfritsch@CLEMSON.EDU

## FEATURES

## source

Location/Qualifiers

1..611

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="C106A13"

/clone\_lib="tomato flower buds 3-8 mm, Cornell University"

/issue\_type="flower"

/dev\_stage="3-8mm buds"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Tanksley, flower buds and flowers were  
taken from greenhouse plants (4-8 wks old, TA496). They  
were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."

BASE COUNT 214 a 90 c 114 g 193 t

## ORIGIN

## Query Match

Best Local Similarity 8.1%; Score 115; DB 118; Length 611.

Matches 329; Conservative 0; Mismatches 235; Indels 33; Gaps 4.

```

UY 80 gaagtggaagagatgctgctacacagggggaagaaagatgctacacaaa 139
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 36 GAAGTGGAAGAGATGCTGCTACACAGGGGGAAGAAAGATGCTACACAAA 95
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 140 ctctcttcacacagatgagctcaatgacacagcagcgaatgca 199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 96 TCTCTTCTTCACACAGATGAGCTCAATGACACAGCAGCTGCAATGAC 155
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 200 aactctctcttcacacagatgagctcaatgacacagcagcgaatgca 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 156 TGAATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 212
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 260 agcagatgcaacacatgacatgacatgacatgacatgacatgacatg 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 213 TTCTGACGTCGAACCTTTTCTGACGTCGAACCTTTTCTGACGTCGA 272
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 420 caaggaatgaattgcaacacatgacatgacatgacatgacatgacatg 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 273 AAAAAACGCACTGTAATGACGCACTGTAATGTAATGTAATGTAATG 432
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 480 tgaattgaatgacatgacatgacatgacatgacatgacatgacatg 432
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 433 TGATTTAAAGCATTTTGTGATGTTAGCGGAAATTGACAAATTTGAAA 392
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 433 gtaagaaatgacatgacatgacatgacatgacatgacatgacatg 487
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 493 TGCAGAACATGCTGACATGCTTTTCTGACGTCGAACCTTTGACGTC 452
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 488 ttttctcttcacacagatgagctcaatgacacagcagcgaatgca 547
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 453 TTTCTCATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTA 512
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 548 ggaagaaatgacatgacatgacatgacatgacatgacatgacatg 607
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 513 AGTTTAAATTTAAATGAAAG-----AACCAAGGCAATTAATTAAT 554
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 608 ataaagaaagccttcctgctgctgctgctgctgctgctgctgctg 654
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

```

DB 555 GCGCACTACAGCTCCAGCACTGTAATGAAAGCAATTAATGAAAGCA 611

## RESULTS

AW928688

LOCUS

DEFINITION

AW928688 611 bp mRNA EST 40-MAY-2000  
EST337476 tomato flower buds 8 mm to pre-anthesis, Cornell  
University Lycopersicon esculentum cDNA clone: cTc296.5, mRNA  
sequence.

ACCESSION

AW928688

VERSION

AW928688.1

KEYWORDS

EST.

SOURCE

ORGANISM

REFERENCE

ATTNERS

TITLE

JOURNAL

COMMENT

van der Hoeven, R.S., Bezzeredes, J.L., Matero, A.L., Holt, J.E., Jiang  
F., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Koning  
C.M., Fraser, C.M., Martin, G.B., Giovannoni, J.J., and Tanksley, S.D.  
Generation of ESTs from tomato flower tissue  
Unpublished (1999)  
Contact: David Fritsch  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 4366  
Fax: 864 656 4293  
Email: dfritsch@CLEMSON.EDU

## FEATURES

source

Location/Qualifiers

1..611

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db\_xref="taxon:4081"

/clone="C106A13"

/clone\_lib="tomato flower buds 8 mm to pre-anthesis,  
Cornell University"

/issue\_type="flower"

/dev\_stage="buds 8mm to pre-anthesis"

/note="Vector: pBluescript SK(+); Site\_1: EcoRI; Site\_2:  
XhoI; supplier: Tanksley, flower buds and flowers were  
taken from greenhouse plants (4-8 wks old, TA496). They  
were immediately frozen in liquid nitrogen and then  
size-separated while remaining frozen."

BASE COUNT 214 a 90 c 114 g 193 t

## ORIGIN

Query Match 8.1%; Score 115; DB 122; Length 611;

Best Local Similarity 55.1%; Pred. No. 1e-16;

Matches 329; Conservative 0; Mismatches 235; Indels 33; Gaps 4;

```

UY 80 gaagtggaagagatgctgctacacagggggaagaaagatgctacacaaa 139
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 36 GAAGTGGAAGAGATGCTGCTACACAGGGGGAAGAAAGATGCTACACAAA 95
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 140 ctctcttcacacagatgagctcaatgacacagcagcgaatgca 199
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 96 TCTCTTCTTCACACAGATGAGCTCAATGACACAGCAGCTGCAATGAC 155
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 200 aactctctcttcacacagatgagctcaatgacacagcagcgaatgca 259
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 156 TGAATGCTATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 212
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 260 agcagatgcaacacatgacatgacatgacatgacatgacatgacatg 319
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
DB 213 TTCTGACGTCGAACCTTTTCTGACGTCGAACCTTTTCTGACGTCGA 272
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||
UY 420 caaggaatgaattgcaacacatgacatgacatgacatgacatgacatg 379
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||

```











Sat Sep 11:45:40 2001

us-09-577-657a-2.rst

Page 13

Job time: 1964 sec

---

...

GenCore version 4.5  
Copyright (c) 1993-2000 CompuGen Ltd.

On multiple - multiple search, using sw model

Run on: August 31, 2001, 13:28:02, Search time 11:38 Seconds

Title: US-09-577-h57A-2

Sequence: 1 tgatatcactgtgtgacag . . . . . attctgggcgcgcgaattc 1427

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 730161 reqs. 414450000 responses

Total number of hits satisfying criterion for method: 146/212

Minimum DB seq length: 0	
Maximum DB seq length: 200000000	

Maximum chr seq length: 2000000000

Post-processing: Minimum Match 0%

Database : N\_Geneseq\_0601:★

1	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1985.DAT
2	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1986.DAT
3	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1987.DAT
4	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1988.DAT
5	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1989.DAT
6	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1990.DAT
7	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1991.DAT
8	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1992.DAT
9	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1993.DAT
10	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1994.DAT
11	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1995.DAT
12	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1996.DAT
13	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1997.DAT
14	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1998.DAT
15	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
16	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
17	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
18	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
19	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
20	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
21	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT
22	/S/SL1/gcg3ata/geneseq/geneseq/nl/NA1999.DAT

pred No is the number of results predicted by chance to have a score greater than or equal to the score or the result being printed and is derived by analysis of the total score distribution.

## SUMMARY

Result No.	Score	Query Match	Length	FR	ID	Description
1	1427	100.0	1427	22	AA082742	C. sinensis N.meth
2	1427	100.0	1427	22	AA082743	C. sinensis N.meth
3	170.4	11.9	1239	21	AA045306	Arabidopsis thaliana
4	170.4	11.9	1457	21	AA045421	Arabidopsis thaliana
5	152.8	10.7	1345	21	AA039444	Arabidopsis thaliana
6	130	9.1	936	22	AA058255	Oligonucleotide D1
7	130	9.1	936	22	AA058254	Oligonucleotide D1
8	130	9.1	936	22	AA058257	Oligonucleotide D1
9	130	9.1	936	22	AA058259	Oligonucleotide D1
10	130	9.1	936	22	AA058262	Oligonucleotide D1
11	130	9.1	938	22	AA058255	Oligonucleotide D1

C	12	1.59	9.0	936	22	AAFS8252	Glucuronidectide D1
C	13	1.59	9.0	936	22	AAFS8254	Glucuronidectide D1
C	14	1.59	9.0	936	22	AAFS8257	Glucuronidectide D1
C	15	1.59	9.0	936	22	AAFS8259	Glucuronidectide D1
C	16	1.59	9.0	936	22	AAFS8262	Glucuronidectide D2
C	17	1.59	9.0	938	22	AAFS8265	Glucuronidectide D1
C	18	72.18	5.11	1211	21	AAFC3567	Arabidopsis thaliana
C	19	72.2	5.1	123	21	AAFC4628	Arabidopsis thaliana
C	20	71.2	5.0	1206	21	AAFC46773	Arabidopsis thaliana
C	21	71.2	5.0	1259	21	AAFC48777	Arabidopsis thaliana
C	22	69	4.8	1261	21	AAFC36551	Arabidopsis thaliana
C	23	65.4	4.7	1273	21	AAFC36551	Arabidopsis thaliana
C	24	65.4	3.7	1244	22	AAFC3238	Arabidopsis thaliana
C	25	62.8	3.7	1122	19	AAVA4616	Human secreted protein
C	26	52.4	3.7	124	22	AAFS8238	Glucuronidectide D1
C	27	52.4	3.7	1281	18	AAFS8749	Pat neuropeptide Y
C	28	52.4	3.7	2181	20	AAVB81195	Pat neuropeptide Y
C	29	52.4	3.7	2604	18	AAAT73595	Pat neuropeptide Y
C	30	52.4	3.7	2604	20	AAVB81196	Pat neuropeptide Y
C	31	52.2	3.7	454	21	AAFC37113	Arabidopsis thaliana
C	32	52.2	3.7	1409	20	AAV3288	Scyban hexose car
C	33	52.2	3.7	2057	21	AAFC78876	Human cancer asso
C	34	51.2	3.6	1282	21	AAFC08972	Fusaria necrotum
C	35	50.4	3.5	1385	21	AAFO8584	Fusarium venenatum
C	36	50.4	3.5	2636	21	AAFO59534	Human secreted protein
C	37	50.4	3.5	2636	21	AAFC64577	Human secreted protein
C	38	50.4	3.5	2636	21	AAFC64577	Human secreted protein
C	39	50.4	3.5	2636	22	AAFA24180	Human secreted protein
C	40	50.4	3.5	2880	21	AAFA24348	cDNA encoding human
C	41	50	3.5	1711	19	AAVA3617	Human secreted protein
C	42	50	3.5	1554	21	AAVA3441	cDNA encoding human
C	43	49.8	3.5	1493	21	AAV32131	Human secreted protein
C	44	49.4	3.5	1410	21	AAV32131	Human secreted protein
C	45	49.4	3.5	2294	22	AAFA5127	Human TANGO-2/3 cD

## ALIGNMENTS

XX	RESULT	1
XX	AAC82742	
XX	AA092742 standard CMA; 1497 RP	
XX	AAC82742;	
XX	16-MAY-2001 (first entry)	
XX	Caffeine biosynthesis, N-methyl transferase cma.	
XX	Caffeine biosynthesis, N-methyl transferase; tea, N-met. plant; ds.	
XX	Camellia sinensis.	
XX	EPI055727-A2.	
XX	29-NOV-2000.	
XX	26 MAY 2000; 2000EP 0304523.	
XX	26-MAY-1999; 99JP-0146358.	
XX	(MITA ) MITSUI CHEM INC.	
XX	Mizuno M, Ashihara H, Mizuno K, Fujimura T;	
XX	WET; 2001 064619708.	
XX	P-PDB; AAB45799.	
XX	CMA encoding N-methyl transferase with enzyme activities of 7 methyl	
XX	transferase, N-methyl transferase, theobromine N-methyl transferase and	
XX	paraxanthine N-methyl transferase, useful for caffeine synthesis in	
XX	plants or microorganisms .	





```

FD 06-SEP-2000.
XX
XX 25-FEB-2000: 2000P 0101449
PR 25-FEB-1999: 990S-0121825.
PR 05-MAR-1999: 990S-0123180.
PR 09-MAR-1999: 990S-0123548.
PR 23-MAR-1999: 990S-0125788.
PR 25-MAR-1999: 990S-0126264.
PR 29-MAR-1999: 990S-0126785.
PR 01-APR-1999: 990S-0127462.
PR 06-APR-1999: 990S-0128234.
PR 08-APR-1999: 990S-0128714.
PR 16-APR-1999: 990S-0129845.
PR 19-APR-1999: 990S-0130077.
PR 21-APR-1999: 990S-0130449.
PR 23-APR-1999: 990S-0130810.
PR 28-APR-1999: 990S-0130891.
PR 28-APR-1999: 990S-0131449.
PR 30-APR-1999: 990S-0132048.
PR 30-APR-1999: 990S-0132407.
PR 04-MAY-1999: 990S-0132484.
PR 05-MAY-1999: 990S-0132485.
PR 06-MAY-1999: 990S-0132486.
PR 06-MAY-1999: 990S-0132487.
PR 07-MAY-1999: 990S-0132863.
PR 11-MAY-1999: 990S-0134256.
PR 14-MAY-1999: 990S-0134218.
PR 14-MAY-1999: 990S-0134219.
PR 14-MAY-1999: 990S-0134221.
PR 14-MAY-1999: 990S-0134370.
PR 18-MAY-1999: 990S-0134768.
PR 19-MAY-1999: 990S-0134941.
PR 20-MAY-1999: 990S-0135124.
PR 21-MAY-1999: 990S-0135353.
PR 24-MAY-1999: 990S-0135629.
PR 25-MAY-1999: 990S-0136021.
PR 27-MAY-1999: 990S-0136392.
PR 28-MAY-1999: 990S-0136782.
PR 01-JUN-1999: 990S-0137222.
PR 03-JUN-1999: 990S-0137228.
PR 04-JUN-1999: 990S-0137520.
PR 07-JUN-1999: 990S-0137724.
PR 08-JUN-1999: 990S-0138094.
PR 10-JUN-1999: 990S-0138540.
PR 10-JUN-1999: 990S-0138847.
PR 16-JUN-1999: 990S-0139119.
PR 16-JUN-1999: 990S-0139452.
PR 16-JUN-1999: 990S-0139453.
PR 17-JUN-1999: 990S-0139492.
PR 18-JUN-1999: 990S-0139454.
PR 18-JUN-1999: 990S-0139455.
PR 18-JUN-1999: 990S-0139456.
PR 18-JUN-1999: 990S-0139457.
PR 18-JUN-1999: 990S-0139458.
PR 18-JUN-1999: 990S-0139459.
PR 18-JUN-1999: 990S-0139460.
PR 18-JUN-1999: 990S-0139461.
PR 18-JUN-1999: 990S-0139462.
PR 18-JUN-1999: 990S-0139463.
PR 18-JUN-1999: 990S-0139750.
PR 18-JUN-1999: 990S-0139763.
PR 21-JUN-1999: 990S-0139817.
PR 22-JUN-1999: 990S-0139899.
PR 23-JUN-1999: 990S-0140353.
PR 23-JUN-1999: 990S-0140354.
PR 24-JUN-1999: 990S-0140695.
PR 28-JUN-1999: 990S-0140823.
PR 29-JUN-1999: 990S-0140991.
PR 30-JUN-1999: 990S-0141287.
PR 01-JUL-1999: 990S-0141842.
PR 01-JUL-1999: 990S-0142154.
PR 02-JUL-1999: 990S-0142055.
PR 06-JUL-1999: 990S-0142300.
PR 08-JUL-1999: 990S-0142803.
PR 09-JUL-1999: 990S-0142920.
PR 12-JUL-1999: 990S-0142977.
PR 13-JUL-1999: 990S-0143542.
PR 14-JUL-1999: 990S-0143624.
PR 15-JUL-1999: 990S-0144005.
PR 16-JUL-1999: 990S-0144085.
PR 16-JUL-1999: 990S-0144086.
PR 16-JUL-1999: 990S-0144125.
PR 19-JUL-1999: 990S-0144131.
PR 19-JUL-1999: 990S-0144132.
PR 19-JUL-1999: 990S-0144133.
PR 19-JUL-1999: 990S-0144134.
PR 19-JUL-1999: 990S-0144334.
PR 20-JUL-1999: 990S-0144335.
PR 20-JUL-1999: 990S-0144352.
PR 20-JUL-1999: 990S-0144463.
PR 20-JUL-1999: 990S-0144484.
PR 21-JUL-1999: 990S-0144414.
PR 21-JUL-1999: 990S-0145086.
PR 21-JUL-1999: 990S-0145088.
PR 22-JUL-1999: 990S-0145085.
PR 22-JUL-1999: 990S-0145087.
PR 22-JUL-1999: 990S-0145089.
PR 22-JUL-1999: 990S-0145192.
PR 23-JUL-1999: 990S-0145145.
PR 23-JUL-1999: 990S-0145218.
PR 23-JUL-1999: 990S-0145224.
PR 25-JUL-1999: 990S-0145276.
PR 25-JUL-1999: 990S-0145277.
PR 27-JUL-1999: 990S-0145918.
PR 27-JUL-1999: 990S-0145919.
PR 28-JUL-1999: 990S-0145951.
PR 02-AUG-1999: 990S-0146186.
PR 02-AUG-1999: 990S-0146188.
PR 02-AUG-1999: 990S-0146189.
PR 03-AUG-1999: 990S-0147048.
PR 04-AUG-1999: 990S-0147204.
PR 04-AUG-1999: 990S-0147402.
PR 05-AUG-1999: 990S-0147132.
PR 05-AUG-1999: 990S-0147460.
PR 06-AUG-1999: 990S-0147303.
PR 06-AUG-1999: 990S-0147416.
PR 06-AUG-1999: 990S-0147493.
PR 06-AUG-1999: 990S-0147935.
PR 10-AUG-1999: 990S-0148171.
PR 11-AUG-1999: 990S-0148319.
PR 12-AUG-1999: 990S-0148341.
PR 13-AUG-1999: 990S-0148565.
PR 13-AUG-1999: 990S-0148684.
PR 16-AUG-1999: 990S-0148684.
PR 17-AUG-1999: 990S-0149175.
PR 18-AUG-1999: 990S-0149426.
PR 20-AUG-1999: 990S-0149722.
PR 20-AUG-1999: 990S-0149723.
PR 20-AUG-1999: 990S-0149924.
PR 20-AUG-1999: 990S-0149925.
PR 20-AUG-1999: 990S-0149930.
PR 23-AUG-1999: 990S-0150666.
PR 23-AUG-1999: 990S-0151080.
PR 30-AUG-1999: 990S-0151303.
PR 31-AUG-1999: 990S-0151438.
PR 01-SEP-1999: 990S-0151930.
PR 07-SEP-1999: 990S-0152363.
PR 10-SEP-1999: 990S-0153070.
PR 13-SEP-1999: 990S-0153758.
PR 15-SEP-1999: 990S-0154018.
PR 16-SEP-1999: 990S-0154039.
PR 20-SEP-1999: 990S-0154779.
PR 22-SEP-1999: 990S-0155139.

```



```

PR 23-SEP-1999: 990S-0156486.
PR 24-SEP-1999: 990S-0156589.
PR 24-SEP-1999: 990S-0156478.
PR 29-SEP-1999: 990S-0156596.
PR 04-OCT-1999: 990S-0157117.
PR 05-OCT-1999: 990S-0157753.
PR 06-OCT-1999: 990S-0157865.
PR 07-OCT-1999: 990S-0158029.
PR 08-OCT-1999: 990S-0158232.
PR 12-OCT-1999: 990S-0158369.
PR 13-OCT-1999: 990S-0159293.
PR 13-OCT-1999: 990S-0159294.
PR 13-OCT-1999: 990S-0159295.
PR 14-OCT-1999: 990S-0159329.
PR 14-OCT-1999: 990S-0159430.
PR 14-OCT-1999: 990S-0159331.
PR 14-OCT-1999: 990S-0159637.
PR 14-OCT-1999: 990S-0159638.
PR 18-OCT-1999: 990S-0159684.
PR 21-OCT-1999: 990S-0160741.
PR 21-OCT-1999: 990S-0160767.
PR 21-OCT-1999: 990S-0160779.
PR 21-OCT-1999: 990S-0160814.
PR 21-OCT-1999: 990S-0160815.
PR 22-OCT-1999: 990S-0160980.
PR 22-OCT-1999: 990S-0160981.
PR 22-OCT-1999: 990S-0160989.
PR 22-OCT-1999: 990S-0161034.
PR 25-OCT-1999: 990S-0161405.
PR 25-OCT-1999: 990S-0161406.
PR 25-OCT-1999: 990S-0161359.
PR 25-OCT-1999: 990S-0161360.
PR 26-OCT-1999: 990S-0161361.
PR 28-OCT-1999: 990S-0161329.
PR 28-OCT-1999: 990S-0161992.
PR 28-OCT-1999: 990S-0161993.
PR 29-OCT-1999: 990S-0162142.

```

Query Match: 11 98; Score: 170.4; PR: 21; Length: 1239.

Best Local Similarity: 46.78; Prod. No: 2.1e-26; Matches: 534; Conservative: 0; Mismatches: 516; Indels: 24; Gaps: 3;

```

QY 91 aagtggttcacacagagggaggaaggaagtagtatgcacaaactcttcttca 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 aaggtctatgataagagtgagtgagtgagtgagtgagtgagtgagtgag 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 151 cgcacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 agaaaagagtgatatacagtgacaaacagtgatgagagagagagagagag 300
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 211 cagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 270
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 tgaacttgaacttcttcttcttcttcttcttcttcttcttcttcttcttct 360
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 271 acacattgcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 330
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 aactcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttct 420
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 331 atgacacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 390
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 421 acaaaacacacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 480
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 391 cgccttcaagagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 447
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 481 cgcacttcaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtg 540
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 gtatagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 507
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 541 gtttcgcatagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 600
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 atttatttatttatttatttatttatttatttatttatttatttatttatttatt 567
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

DB 601 atttaaacattctctctctctctctctctctctctctctctctctctctctct 660
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 568 ggcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 627
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 661 atttatttatttatttatttatttatttatttatttatttatttatttatttatt 702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 628 tcttctctctctctctctctctctctctctctctctctctctctctctctct 687
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 703 gtcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 762
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 688 gacccacagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 747
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 763 gtcagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 822
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 748 ctctctctctctctctctctctctctctctctctctctctctctctctctct 804
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 823 ttaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 882
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 805 ctgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 864
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 884 gtagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 942
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 865 gctatttctctctctctctctctctctctctctctctctctctctctctctct 924
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 943 ttaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1002
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 925 ttaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 984
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1003 ttaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1062
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 985 ttaagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1044
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1063 aagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1045 agtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1104
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1123 atttctctctctctctctctctctctctctctctctctctctctctctctct 1182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1105 atttctctctctctctctctctctctctctctctctctctctctctctctct 1158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1183 atttctctctctctctctctctctctctctctctctctctctctctctctct 1246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

# RESULT 4

AAC35921  
ID AAC35921 standard. PNM, 1407 BP

XX AAC35921:  
XX

CT 17-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 11896.

XX Hybridisation assay, genetic mapping, gene expression control;  
KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.

XX 05-MAR-1999; 990S-0123180.

XX 06-MAR-1999; 990S-0123648.

XX 23-MAR-1999; 990S-0126788.

XX 25-MAR-1999; 990S-0126785.

XX 06-APR-1999; 990S-0127462.

08-APR-1999: 99US-0128714.  
PR 16-APR-1999: 99US-0129845.  
PR 19-APR-1999: 99US-0130077.  
PR 21-APR-1999: 99US-0130449.  
PR 23-APR-1999: 99US-0130510.  
PR 24-APR-1999: 99US-0130891.  
PR 28-APR-1999: 99US-0131449.  
PR 30-APR-1999: 99US-0132048.  
PR 04-APR-1999: 99US-0132407.  
PR 04-MAY-1999: 99US-0132484.  
PR 06-MAY-1999: 99US-0132485.  
PR 06-MAY-1999: 99US-0132486.  
PR 07-MAY-1999: 99US-0132487.  
PR 11-MAY-1999: 99US-0132863.  
PR 14-MAY-1999: 99US-0134256.  
PR 14-MAY-1999: 99US-0134218.  
PR 14-MAY-1999: 99US-0134219.  
PR 14-MAY-1999: 99US-0134221.  
PR 18-MAY-1999: 99US-0134370.  
PR 19-MAY-1999: 99US-0134768.  
PR 20-MAY-1999: 99US-0134941.  
PR 21-MAY-1999: 99US-0135353.  
PR 24-MAY-1999: 99US-0135629.  
PR 25-MAY-1999: 99US-0136021.  
PR 27-MAY-1999: 99US-0136392.  
PR 28-MAY-1999: 99US-0136782.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 07-JUN-1999: 99US-0137724.  
PR 08-JUN-1999: 99US-0138094.  
PR 10-JUN-1999: 99US-0138540.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 17-JUN-1999: 99US-0139492.  
PR 18-JUN-1999: 99US-0139454.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139763.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139899.  
PR 23-JUN-1999: 99US-0140153.  
PR 23-JUN-1999: 99US-0140354.  
PR 23-JUN-1999: 99US-0140355.  
PR 28-JUN-1999: 99US-0140823.  
PR 28-JUN-1999: 99US-0140961.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142055.  
PR 06-JUL-1999: 99US-0142366.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142820.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 14-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.  
PR 16-JUL-1999: 99US-0144085.  
PR 16-JUL-1999: 99US-0144086.  
PR 19-JUL-1999: 99US-0144325.  
PR 19-JUL-1999: 99US-0144331.  
PR 19-JUL-1999: 99US-0144332.  
  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144335.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144432.  
PR 20-JUL-1999: 99US-0144484.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0144896.  
PR 21-JUL-1999: 99US-0145088.  
PR 22-JUL-1999: 99US-0145085.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145224.  
PR 26-JUL-1999: 99US-0145276.  
PR 27-JUL-1999: 99US-0145918.  
PR 27-JUL-1999: 99US-0145919.  
PR 28-JUL-1999: 99US-0145951.  
PR 02-AUG-1999: 99US-0146386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146389.  
PR 03-AUG-1999: 99US-0147038.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147202.  
PR 07-AUG-1999: 99US-0147392.  
PR 07-AUG-1999: 99US-0147360.  
PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147494.  
PR 10-AUG-1999: 99US-0148171.  
PR 11-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148341.  
PR 13-AUG-1999: 99US-0148365.  
PR 13-AUG-1999: 99US-0148684.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149175.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149723.  
PR 20-AUG-1999: 99US-0149924.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149930.  
PR 25-AUG-1999: 99US-0150566.  
PR 26-AUG-1999: 99US-0150884.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151066.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151403.  
PR 31-AUG-1999: 99US-0151438.  
PR 01-SEP-1999: 99US-0151930.  
PR 07-SEP-1999: 99US-0152363.  
PR 10-SEP-1999: 99US-0153070.  
PR 13-SEP-1999: 99US-0153758.  
PR 15-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 16-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155654.  
PR 28-SEP-1999: 99US-0156458.  
PR 29-SEP-1999: 99US-0156506.  
PR 04-OCT-1999: 99US-0157117.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158232.  
PR 12-OCT-1999: 99US-0158369.  
PR 13-OCT-1999: 99US-0159293.  
PR 13-OCT-1999: 99US-0159294.



PK 06-MAY-1999; 9905-0132487.  
PK 07-MAY-1999; 9905-0132863.  
PK 11-MAY-1999; 9905-0134256.  
PK 14-MAY-1999; 9905-0134218.  
PK 14-MAY-1999; 9905-0134219.  
PK 14-MAY-1999; 9905-0134221.  
PK 14-MAY-1999; 9905-0134370.  
PK 14-MAY-1999; 9905-0134768.  
PK 14-MAY-1999; 9905-0134941.  
PK 20-MAY-1999; 9905-0135124.  
PK 21-MAY-1999; 9905-0135353.  
PK 24-MAY-1999; 9905-0135629.  
PK 25-MAY-1999; 9905-0136021.  
PK 27-MAY-1999; 9905-0136392.  
PK 28-MAY-1999; 9905-0136782.  
PK 01-JUN-1999; 9905-0137222.  
PK 03-JUN-1999; 9905-0137528.  
PK 04-JUN-1999; 9905-0137502.  
PK 07-JUN-1999; 9905-0137724.  
PK 08-JUN-1999; 9905-0138094.  
PK 10-JUN-1999; 9905-0138540.  
PK 10-JUN-1999; 9905-0138847.  
PK 14-JUN-1999; 9905-0139119.  
PK 16-JUN-1999; 9905-0139452.  
PK 16-JUN-1999; 9905-0139453.  
PK 17-JUN-1999; 9905-0139492.  
PK 18-JUN-1999; 9905-0139454.  
PK 18-JUN-1999; 9905-0139455.  
PK 18-JUN-1999; 9905-0139456.  
PK 18-JUN-1999; 9905-0139457.  
PK 18-JUN-1999; 9905-0139459.  
PK 18-JUN-1999; 9905-0139459.  
PK 18-JUN-1999; 9905-0139460.  
PK 18-JUN-1999; 9905-0139461.  
PK 18-JUN-1999; 9905-0139462.  
PK 18-JUN-1999; 9905-0139463.  
PK 18-JUN-1999; 9905-0139750.  
PK 18-JUN-1999; 9905-0139763.  
PK 21-JUN-1999; 9905-0139817.  
PK 22-JUN-1999; 9905-0139899.  
PK 23-JUN-1999; 9905-0140353.  
PK 23-JUN-1999; 9905-0140354.  
PK 24-JUN-1999; 9905-0140354.  
PK 24-JUN-1999; 9905-0140695.  
PK 28-JUN-1999; 9905-0140823.  
PK 29-JUN-1999; 9905-0140991.  
PK 30-JUN-1999; 9905-0141287.  
PK 01-JUL-1999; 9905-0141842.  
PK 01-JUL-1999; 9905-0142154.  
PK 02-JUL-1999; 9905-0142055.  
PK 06-JUL-1999; 9905-0142390.  
PK 08-JUL-1999; 9905-0142803.  
PK 09-JUL-1999; 9905-0142920.  
PK 12-JUL-1999; 9905-0142977.  
PK 13-JUL-1999; 9905-0143542.  
PK 14-JUL-1999; 9905-0143624.  
PK 15-JUL-1999; 9905-0144005.  
PK 16-JUL-1999; 9905-0144086.  
PK 16-JUL-1999; 9905-0144086.  
PK 19-JUL-1999; 9905-0144325.  
PK 19-JUL-1999; 9905-0144331.  
PK 19-JUL-1999; 9905-0144332.  
PK 19-JUL-1999; 9905-0144333.  
PK 19-JUL-1999; 9905-0144334.  
PK 19-JUL-1999; 9905-0144335.  
PK 20-JUL-1999; 9905-0144352.  
PK 20-JUL-1999; 9905-0144632.  
PK 20-JUL-1999; 9905-0144684.  
PK 21-JUL-1999; 9905-0144814.  
PK 21-JUL-1999; 9905-0145086.  
PK 21-JUL-1999; 9905-0145088.  
PK 22-JUL-1999; 9905-0145085.  
PK 22-JUL-1999; 9905-0145087.  
PK 22-JUL-1999; 9905-0145089.  
  
PK 22-JUL-1999; 9905-0145192.  
PK 23-JUL-1999; 9905-0145145.  
PK 23-JUL-1999; 9905-0145218.  
PK 23-JUL-1999; 9905-0145224.  
PK 26-JUL-1999; 9905-0145276.  
PK 27-JUL-1999; 9905-0145913.  
PK 27-JUL-1999; 9905-0145918.  
PK 27-JUL-1999; 9905-0145919.  
PK 28-JUL-1999; 9905-0145951.  
PK 02-AUG-1999; 9905-0146386.  
PK 02-AUG-1999; 9905-0146388.  
PK 03-AUG-1999; 9905-0146389.  
PK 03-AUG-1999; 9905-0147038.  
PK 04-AUG-1999; 9905-0147204.  
PK 04-AUG-1999; 9905-0147302.  
PK 05-AUG-1999; 9905-0147192.  
PK 06-AUG-1999; 9905-0147260.  
PK 06-AUG-1999; 9905-0147303.  
PK 06-AUG-1999; 9905-0147416.  
PK 09-AUG-1999; 9905-0147493.  
PK 09-AUG-1999; 9905-0147935.  
PK 10-AUG-1999; 9905-0148171.  
PK 11-AUG-1999; 9905-0148319.  
PK 12-AUG-1999; 9905-0148341.  
PK 13-AUG-1999; 9905-0148565.  
PK 13-AUG-1999; 9905-0148684.  
PK 16-AUG-1999; 9905-0149368.  
PK 17-AUG-1999; 9905-0149175.  
PK 18-AUG-1999; 9905-0149426.  
PK 20-AUG-1999; 9905-0149722.  
PK 20-AUG-1999; 9905-0149723.  
PK 20-AUG-1999; 9905-0149929.  
PK 23-AUG-1999; 9905-0149932.  
PK 23-AUG-1999; 9905-0149930.  
PK 25-AUG-1999; 9905-0150506.  
PK 26-AUG-1999; 9905-0150884.  
PK 27-AUG-1999; 9905-0151065.  
PK 27-AUG-1999; 9905-0151066.  
PK 27-AUG-1999; 9905-0151080.  
PK 30-AUG-1999; 9905-0151303.  
PK 31-AUG-1999; 9905-0151303.  
PK 01-SEP-1999; 9905-0151930.  
PK 07-SEP-1999; 9905-0152163.  
PK 10-SEP-1999; 9905-0153070.  
PK 13-SEP-1999; 9905-0153758.  
PK 15-SEP-1999; 9905-0154018.  
PK 16-SEP-1999; 9905-0154039.  
PK 20-SEP-1999; 9905-0154779.  
PK 22-SEP-1999; 9905-0155139.  
PK 23-SEP-1999; 9905-0155486.  
PK 24-SEP-1999; 9905-0155664.  
PK 28-SEP-1999; 9905-0156458.  
PK 29-SEP-1999; 9905-0156596.  
PK 04-OCT-1999; 9905-0157117.  
PK 05-OCT-1999; 9905-0157753.  
PK 06-OCT-1999; 9905-0157865.  
PK 07-OCT-1999; 9905-0158229.  
PK 08-OCT-1999; 9905-0158332.  
PK 12-OCT-1999; 9905-0158369.  
PK 13-OCT-1999; 9905-0159293.  
PK 13-OCT-1999; 9905-0159294.  
PK 13-OCT-1999; 9905-0159295.  
PK 14-OCT-1999; 9905-0159329.  
PK 14-OCT-1999; 9905-0159330.  
PK 14-OCT-1999; 9905-0159331.  
PK 14-OCT-1999; 9905-0159637.  
PK 14-OCT-1999; 9905-0159638.  
PK 18-OCT-1999; 9905-0163584.  
PK 21-OCT-1999; 9905-0160741.  
PK 21-OCT-1999; 9905-0160767.  
PK 21-OCT-1999; 9905-0160768.  
PK 21-OCT-1999; 9905-0160770.  
PK 21-OCT-1999; 9905-0160814.



```

UY 628 ttatagaagaagctactatctcaattcaatgaagatttcgaatgtttccaaatgcta 687
IB 2 www. .... 61
UY 688 gatcccaagagagatcccaatgattatgatttgaactcgtgaagcaatglt 747
IB 62 www. .... 121
UY 748 ctatctcttcgaatcgaatgcttgaactcgaatcattatcgaatgcaatgctg 807
IB 122 www. .... 181
UY 808 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 867
IB 182 www. .... 241
UY 868 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 927
IB 242 www. .... 301
UY 928 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 987
IB 402 www. .... 361
UY 988 gatctgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1047
IB 462 www. .... 421
UY 1048 ttggaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1107
IB 422 www. .... 481
UY 1108 ttggaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1167
IB 482 www. .... 541
UY 1168 atttttgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1227
IB 542 www. .... 601
UY 1228 gttgaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1287
IB 602 www. .... 661
UY 1288 ttggaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1347
IB 662 www. .... 721
UY 1348 atttttgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1407
IB 722 www. .... 781
UY 1408 atttttgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1467
IB 782 www. .... 841

```

```

XX 01-FEB-2001.
XX
XX 26-JUL-2000; 2009MO-US20476.
XX
XX 26-JUL-1999; 592S-G145695.
XX 17-MAR-2000; 2009S-0190259.
XX
XX (CLIN) CLINICAL MICRO SENSORS INC.
XX
XX time RM:
XX
XX WFL, 2001 1597:28/16.
XX
XX Nucleic acids containing electron transfer group, useful as labels in
XX hybridization assays, e.g. for genotyping, allowing repeat analyses on
XX a single surface.
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
XX acids each containing an electron transfer group (ETG) having
XX different redox potentials. The invention is used for electronic
XX detection of nucleic acids, especially of substitutions (mismatches)
XX and single nucleotide polymorphisms, e.g. for genotyping.
XX monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
XX
Query Match 9.1%; Score 130; DB 22; Length 946;
Best Local Similarity 1.3%; Pred. No. 2,5e-20;
Matches 10; Conservative 493; Mismatches 292; Indels 0; Caps 0;
UY 628 ttgtaagaagaagcctactatcctcaattcagaaatttcaaatgttttcaatgcta 687
DB 2 www. .... 61
UY 688 gatcccaagagatgattcgaatgattgaatgatttgaatcctgtaagatgatt 747
DB 62 www. .... 121
UY 748 ctatctcttcgaatcgaatgcttgaactcgaatcattatcgaatgcaatgctg 807
DB 122 www. .... 181
UY 808 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 867
DB 182 www. .... 241
UY 868 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 927
DB 242 www. .... 301
UY 928 attttgatttcaagatgattgaatgaatgaatgaatgaatgaatgaatgaatgaat 987
DB 302 www. .... 361
UY 988 gatctgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1047
DB 362 www. .... 421
UY 1048 ttggaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1107
DB 422 www. .... 481
UY 1108 ttggaactgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1167
DB 482 www. .... 541
UY 1168 atttttgaatgattgaatgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1227
DB 542 www. .... 601

```

```

RESULT 7
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX
XX AAF58254:
XX
XX 24-APR-2001 (first entry)
XX
XX oligonucleotide D1875.
XX
XX Electron transfer group; ETG; mismatch; genotyping;
XX gene expression; ss.
XX
XX Synthetic.
XX
XX W0200107665-A2.

```

QY 1228 gtagcaatgcatgacacaaatctgagaggttcaaatatataagacatttgcctc 1287  
Db 602 www. .... 661  
QY 1288 tgggtggagagaggtttcttgatttaattctgtgatacccaattcgttatttgg 1347  
Db 662 www. .... 721  
QY 1348 aagaaatgagaaatltgaacatgaattttaaaaaaataaaaaaataaaaaa 1407  
Db 722 www. .... 781  
QY 1408 attctcgagcgccgc 1421  
Db 782 www. .... 795

RESULT 8  
AAF58257  
ID AAF58257 standard: DNA: 936 BP.  
XX  
AC AAF58257:  
XX  
DI 24-APR-2001 (first entry)  
XX  
DB Oligonucleotide D1954.  
XX  
KW Election-transfer group: ETM; mismatch: genotyping.  
KM gene expression: ss.  
XX  
OS Synthetic.  
XX  
PN WO200107665-A2.  
XX  
PD 01-FEB-2001  
XX  
PF 26-JUL-2000: 2000WO-US20476.  
XX  
PR 26-JUL-1999: 99US-0147695.  
PR 17-MAR-2000: 2000US-0190259.  
XX  
PA (CLIN-) CLINICAL MICRO SENSORS INC.  
XX  
PI time RM:  
XX  
DR WPI: 2001-159728/16.  
XX  
PT Nucleic acids containing election-transfer group, useful as labels in  
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
XX a single surface  
XX  
PS Example 6: Page 127: 159pp; English.  
XX  
CC The present invention relates to a composition comprising two nucleic  
CC acids each containing an election-transfer group (ETM) having  
CC different redox potentials. The invention is used for electronic  
CC detection of nucleic acids, especially of substitutions (mismatches)  
CC and single-nucleotide polymorphisms, e.g. for genotyping.  
CC monitoring gene expression.  
XX  
SQ Sequence 936 BP: 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 9.18; Score 130; DB 22; Length 936;  
Best Local Similarity 1.38; Pctd. No. 2.56-20;  
Matches 10; Conservative 492; Mismatches 292; Indels 0; Gaps 0;

QY 628 ttgaaagaagactatcatcatcatcatgaagaatttcaaatgtttccaatgcta 687  
Db 2 www. .... 61  
QY 688 gatcccaagagtgatccaatgattgatatgttgtaactcgtgtgagcaatgtt 747

Db 62 www. .... 121  
QY 748 ctgactctcagacacagagatgcttacttggaactatgactatgttcattgttg 807  
Db 122 www. .... 181  
QY 908 aattggttccagggattgatatgatatgaatatataacattcaataaccagct 867  
Db 182 www. .... 241  
QY 868 atttgcataacttggaggaagtgaagatacatagtgaagggagcagatcacaatg 927  
Db 242 www. .... 301  
QY 928 atcataagagggatttgcattgtagcgtagaatgcagagagaatataatggta 987  
Db 302 www. .... 361  
QY 988 gaggagaaagttaccagagttgcagggttcacagaggtataattccaaccagt 1047  
Db 362 www. .... 421  
QY 1048 ttgacctgaatcatgacaaatataagacaattcattacatgttgatttcagatt 1107  
Db 422 www. .... 481  
QY 1108 tgaagcaagctacagaaaccacaaaglatcatctagtttccaagattgagat 1167  
Db 482 www. .... 541  
QY 1168 agtttlltagtlttgaaataacacgttgcctatcacatatgcacatagaggtt 1227  
Db 542 www. .... 601  
QY 1228 gtagcaatgcatgacacaaatctgagaggttcaaatatataagacatttgcctc 1287  
Db 602 www. .... 661  
QY 1288 tgggtggagagaggtttcttgatttaattctgtgatacccaattcgttatttgg 1347  
Db 662 www. .... 721  
QY 1348 aagaaatgagaaatltgaacatgaattttaaaaaaataaaaaaataaaaaa 1407  
Db 722 www. .... 781  
QY 1408 attctcgagcgccgc 1421  
Db 782 www. .... 795

RESULT 9  
AAF58259  
ID AAF58259 standard: DNA: 936 BP.  
XX  
AC AAF58259:  
XX  
DI 24-APR-2001 (first entry)  
XX  
DB Oligonucleotide D2004.  
XX  
KW Election-transfer group: ETM; mismatch: genotyping.  
KM gene expression: ss.  
XX  
OS Synthetic.  
XX  
PN WO200107665-A2.  
XX  
PD 01-FEB-2001  
XX  
PF 26-JUL-2000: 2000WO-US20476.  
XX





```

Db 122 www..... 181
OY 808 aattggttcaaggagatgagatgaataaattgaacccctcaatataccagct 867
Db 182 www..... 241
OY 868 atttgaatattgagagatgaaagatatattgagagagagatattgagattg 927
Db 242 www..... 301
OY 928 atatatagaggggtttgtttgtatgctgagagagagagagagagagagag 987
Db 302 www..... 361
OY 988 gaaggagaaattacaaagctgcaaggccctccacagacataattcaaccagt 1047
Db 362 www..... 421
OY 1048 ttgacctgaatacaggaacataatagacaattcactcaattgagttcagatt 1107
Db 422 www..... 481
OY 1108 tggagagaaagctacaggaagagagagatgagagagagagagagagagag 1167
Db 482 www..... 541
OY 1168 agtttttttttttttttttttttttttttttttttttttttttttttttt 1227
Db 542 www..... 601
OY 1228 gtgcacatgttttgcagagagagagagagagagagagagagagagagag 1287
Db 602 www..... 661
OY 1288 tttgtgagagagagagagagagagagagagagagagagagagagagag 1347
Db 662 www..... 721
OY 1348 aaggaatgagagagagagagagagagagagagagagagagagagagag 1407
Db 722 www..... 781
OY 1408 attcctgagcgcgc 1421
Db 782 www..... 795

```

```

RESULT 11
AAF58255
ID AAF58255 standard: DNA: 938 BP
XX
AC AAF58255:
XX
DT 24 APR 2001 (first entry)
XX
DE Oligonucleotide D1876.
XX
KM Election-transfer group, ETM, mismatch, genotyping;
KW gene expression, SS.
XX
OS Synthetic.
XX
PN W020010/565-A2.
XX
FE 01 FEB 2001
XX
PE 26-JUL-2000: 2000W0-JS20476.
XX
PR 26-JUL-1999: 94NS-0145645
PR 17-MAR-2000: 2000NS-0140254
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX

```

```

P1 Unk RM:
XX
DR WPI: 2001-159728/16.
XX
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
PS Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electro-
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SO Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other:

Query Match 9.1%; Score 100; PR 22; Length 938;
Best Local Similarity 1.3%; Prod. No. 2.5e-20;
Matches 10; Conserved 10; 40; Mismatches 29; Indels 0; Gaps 0;

OY 628 ttgagagagagagagagagagagagagagagagagagagagagagagagag 687
Db 2 www..... 61
OY 688 gatccaaagagagagagagagagagagagagagagagagagagagagagag 747
Db 62 www..... 121
OY 748 atgacttttgaatgagagagagagagagagagagagagagagagagagag 807
Db 122 www..... 181
OY 808 aattggttcaaggagatgagatgaataaattgaacccctcaatataccagct 867
Db 182 www..... 241
OY 868 atttgaatattgagagatgaaagatatattgagagagagagagagagagag 927
Db 242 www..... 301
OY 928 atatatagaggggtttgtttgtatgctgagagagagagagagagagagag 987
Db 302 www..... 361
OY 988 gaaggagaaattacaaagctgcaaggccctccacagacataattcaaccagt 1047
Db 362 www..... 421
OY 1048 ttgacctgaatacaggaacataatagacaattcactcaattgagttcagatt 1107
Db 422 www..... 481
OY 1108 tggagagaaagctacaggaagagagagagagagagagagagagagagag 1167
Db 482 www..... 541
OY 1168 agtttttttttttttttttttttttttttttttttttttttttttttttt 1227
Db 542 www..... 601
OY 1228 gtgcacatgttttgcagagagagagagagagagagagagagagagagag 1287
Db 602 www..... 661
OY 1288 tttgtgagagagagagagagagagagagagagagagagagagagagag 1347
Db 662 www..... 721
OY 1348 aaggaatgagagagagagagagagagagagagagagagagagagagag 1407

```



XX The present invention relates to a composition comprising two nucleic  
 CC acids each containing an electron-transfer group (ETG) having  
 CC different redox potentials. The invention is used for electroc  
 CC detection of nucleic acids, especially of substitutions (mismatches)  
 CC and single-nucleotide polymorphisms, e.g. for genotyping.  
 CC monitoring gene expression.

XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 9.0%; Score 129; DB 22; Length 936;

Best Local Similarity 0.8%; Pred. No. 4.2e-20;

Matches 6; Conservative 484; Mismatches 279; Indels 0; Gaps 0;

DB 642 taactatcattcatgaagattcacaatglttccaatgtagatcccaagatg 701

DB 779 ww 720

DB 702 gtcccaatggttgatggttgatgacttggtgaggaatggttgatgacttgag 761

DB 719 ww 660

DB 762 atgagagtggttcttgggaactatgtaggagatgagatgagatgagatgag 821

DB 659 ww 600

DB 822 ggaatgataagagaataatagacactcacaataaccagatattgacactt 881

DB 599 ww 540

DB 882 gagagagtgagatgagatgagatgagatgagatgagatgagatgagatgag 941

DB 539 ww 480

DB 942 ttgatctgatagagagagaatgagagagagagagagagagagagagagag 1001

DB 479 ww 420

DB 1002 accaaggttctgagagccttcacagacacacacacacacacacacacacac 1061

DB 419 ww 360

DB 166 atgag 1121

DB 359 ww 300

DB 1122 ccgaagaccacagatcactcactcactcactcactcactcactcactcactc 1181

DB 299 ww 240

DB 1182 gtgagataaactgttgcctcactcactcactcactcactcactcactcactc 1241

DB 249 ww 180

DB 1242 cagcagagagatgagagagagagagagagagagagagagagagagagag 1301

DB 179 ww 120

DB 1302 atgttcttgatttaaatctgtgatacccaatctgataatgttgagagagaat 1361

DB 119 ww 60

DB 1362 tgaacatgaatttcaaaaaaataaaaaaataaaaaaataaaaaaata 1410

DB 59 ww 11

RESULT 14  
 AAF58257/c  
 ID AAF58257 standard; DNA, 936 BP.  
 XX  
 AC AAF58257;

XX 24-Apr-2001 (first entry)  
 XX  
 XX oligonucleotide 11954.

XX  
 XX  
 XX Electron-transfer group; ETG: mismatch; genotyping;  
 KW gene expression; SS.

XX Synthetic.

XX W020010765-A2.

XX  
 XX 01 PER 2001

XX 2e-JUL-2000; 2000W0-US20476.

XX 26-ME-1999; 99NS-0145695.

XX 17-MAR-2000; 2000NS-0190259.

XX (CLIN-) CLINICAL MICRO SENSORS INC.

XX Dmek FM;

XX W01, 2001-15978/15.

XX Nucleic acids containing electron transfer group, useful as labels in  
 PT hybridization assays, e.g. for genotyping, allowing repeat analyses on  
 PT a single surface.

XX Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic  
 CC acids each containing an electron-transfer group (ETG) having  
 CC different redox potentials. The invention is used for electronic  
 CC detection of nucleic acids, especially of substitutions (mismatches)  
 CC and single-nucleotide polymorphisms, e.g. for genotyping,  
 CC monitoring gene expression.

XX Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;

Query Match 9.0%; Score 129; DB 22; Length 936;

Best Local Similarity 0.8%; Pred. No. 4.2e-20;

Matches 6; Conservative 484; Mismatches 279; Indels 0; Gaps 0;

DB 642 taactatcattcatgaagattcacaatglttccaatgtagatcccaagatg 701

DB 779 ww 720

DB 702 gtcccaatggttgatggttgatgacttggtgaggaatggttgatgacttgag 761

DB 719 ww 660

DB 762 atgagagtggttcttgggaactatgtaggagatgagatgagatgagatgag 921

DB 659 ww 600

DB 822 ggaatgataagagaataatagacactcacaataaccagatattgacactt 881

DB 599 ww 540

DB 882 gagagagtgagatgagatgagatgagatgagatgagatgagatgagatgag 941

DB 539 ww 480

DB 942 ttgatctgatagagatgagagatgagagatgagagatgagagatgagagatg 1001

DB 479 ww 420

DB 1002 accaaggttctgagagccttcacagacacacacacacacacacacacacac 1061

DB 419 ww 360



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 31, 2001, 13:25:47, Search time 1922.14 Seconds

(without alignments)  
11483.283 Million cell updates/sec

Title: US-09-577-657A-2

Perfect score: 1427

Sequence: 1 tggatcagcagtgaggag

atttttgggtttgggatttt 1427

Scoring table: IDENTITY\_NUC

Gapop 10.0, Gapext 1.0

Searched: 1344157 seqs, 773874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: GenEmbl:\*

1: qb\_ba1:\*

2: qb\_ba2:\*

3: qb\_ba3:\*

4: qb\_in1:\*

5: qb\_in2:\*

6: qb\_in3:\*

7: qb\_om:\*

8: qb\_ov:\*

9: qb\_pat1:\*

10: qb\_pat2:\*

11: qb\_ph:\*

12: qb\_pl1:\*

13: qb\_pl2:\*

14: qb\_pl3:\*

15: qb\_pl4:\*

16: em\_ba1:\*

17: em\_ba2:\*

18: em\_fun:\*

19: em\_hteo\_hum:\*

20: em\_hteo\_inv:\*

21: em\_hteo\_rod:\*

22: em\_htrg\_hum1:\*

23: em\_htrg\_hum2:\*

24: em\_htrg\_hum3:\*

25: em\_htrg\_hum4:\*

26: em\_htrg\_hum5:\*

27: em\_htrg\_hum6:\*

28: em\_htrg\_hum7:\*

29: em\_htrg\_hum8:\*

30: em\_htrg\_inv1:\*

31: em\_htrg\_inv2:\*

32: em\_htrg\_other:\*

33: em\_htrg\_rod:\*

34: em\_hum1:\*

35: em\_hum2:\*

36: em\_hum3:\*

37: em\_hum4:\*

38: em\_hum5:\*

39: em\_hum6:\*

40: em\_hum7:\*

41: em\_in:\*

42: em\_om:\*

43: em\_ov:\*

44: em\_ov:\*

45: em\_pat:\*

46: em\_ph:\*

47: em\_pl:\*

48: em\_rod:\*

49: em\_sts:\*

50: em\_sy:\*

51: em\_un:\*

52: em\_v1:\*

53: qb\_sts1:\*

54: qb\_sts2:\*

55: qb\_sts3:\*

56: qb\_sy:\*

57: qb\_un:\*

58: qb\_v11:\*

59: qb\_v12:\*

60: qb\_htrg1:\*

61: qb\_htrg2:\*

62: qb\_htrg3:\*

63: qb\_htrg4:\*

64: qb\_htrg5:\*

65: qb\_htrg6:\*

66: qb\_htrg7:\*

67: qb\_htrg8:\*

68: qb\_htrg9:\*

69: qb\_htrg10:\*

70: qb\_htrg11:\*

71: qb\_htrg12:\*

72: qb\_htrg13:\*

73: qb\_htrg14:\*

74: qb\_htrg15:\*

75: qb\_htrg16:\*

76: qb\_htrg17:\*

77: qb\_htrg18:\*

78: qb\_htrg19:\*

79: qb\_htrg20:\*

80: qb\_htrg21:\*

81: qb\_htrg22:\*

82: qb\_htrg23:\*

83: qb\_htrg24:\*

84: qb\_htrg25:\*

85: qb\_pr1:\*

86: qb\_pr2:\*

87: qb\_pr3:\*

88: qb\_pr4:\*

89: qb\_pr5:\*

90: qb\_pr6:\*

91: qb\_pr7:\*

92: qb\_pr8:\*

93: qb\_pr9:\*

94: qb\_pr10:\*

95: qb\_pr11:\*

96: qb\_pr12:\*

97: qb\_pr13:\*

98: em\_ba3:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1403.2	98.4	1408	12	AB041260
2	238.2	16.7	1347	15	AB042752
3	215.4	15.1	1152	15	AF1308570
4	213.8	15.0	1321	12	AF133053
5	155.6	10.9	1306	13	AF176222
6	151.6	10.6	1404	12	AB044702
7	151.2	10.6	1316	12	AB048793
8	149.2	10.5	1359	12	AB039725

9	138.4	9.7	1.29	12	AB048794	Coltea a
10	134.4	9.4	1.33	13	AB198492	AB198492
11	103.8	7.3	93890	12	AC006528	Artichrpn
12	101.4	7.1	82348	12	AB028622	Arabidops
13	100.6	7.0	61510	12	AB028606	Arabidops
14	94.2	6.5	198354	13	ATAP22	279708 Arabidops
15	93.2	6.5	198750	13	ATCHR1V85	AB114589 Arabidops
16	92.8	6.5	83511	12	AB013485	AB013389 Arabidops
17	84	5.9	117206	12	AC008153	AB013389 Arabidops
18	83	5.8	90341	12	AC004609	AC004609 Arabidops
19	79.6	5.6	454	14	BVU234779	AC234779 Arabidops
20	71	5.0	102565	12	AC025563	AC025563 Arabidops
21	69.6	4.9	149477	83	AP003378	AP003378 Oryza sat
22	69.6	4.9	160458	83	AP000345	AP001247 Oryza sat
23	69.2	4.8	151629	83	AP008354	AT031254 Arabidops
24	69.2	4.8	49890	13	ATP28B10	AT141254 Arabidops
25	55.8	3.9	78239	12	AB012241	AB313241 Arabidops
26	54.2	3.8	1084	53	CNS006PL8	AB1409509 17 end of
27	52.6	3.7	3209	5	AT180728	AT180728 Arabidops
28	52.6	3.7	2550	93	HSM801250	AT120428 Arabidops
29	52.6	3.7	124008	87	AC006882	AC006882 Homo sapi
30	52.6	3.7	145264	87	AC016601	AC016601 Homo sapi
31	52.6	3.7	160583	69	AC026006	AC026006 Homo sapi
32	52.6	3.7	164399	96	PMAL33P6	298551 Plasmidium
33	52.6	3.7	2175	95	RMCEC274	065274 Rattus norv
34	52.4	3.7	2481	9	AR079011	AR079011 Sequence
35	52.4	3.7	2604	9	AP079011	AP079012 Sequence
36	52.2	3.6	141964	14	FTPL1	AC007551 Arabidops
37	51.4	3.6	1380	91	BC053373	BC053373 Homo sapi
38	51.4	3.6	49890	13	ATP28B10	AT141254 Arabidops
39	51	3.6	779	53	CNS064XV	AT141257 17 end of
40	51	3.6	47460	12	AB011478	AB011478 Arabidops
41	51	3.6	164446	60	AC008792	AC008792 Homo sapi
42	51	3.6	214094	77	AC087188	AC087188 Homo sapi
43	50.4	3.5	1868	89	AP210052	AP210052 Homo sapi
44	50.2	3.5	2152	5	AY008296	AY008296 Homo sapi
45	50.2	3.5	143132	76	AC026436	AC026436 Homo sapi

## ALIGNMENTS

[illegible]

RESULT	1
LOCUS	AB031280
DEFINITION	AB031280 1438 bp mRNA
ACCESSION	AB031280
VERSION	AB031280.1 GI:9967142
KEYWORDS	caffeine synthase
SOURCE	Camellia sinensis cDNA to mRNA.
ORGANISM	Camellia sinensis Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta Magnoliophyta; Eudicotyledons; Core eudicots; Asteridae; Fricales; Theaceae; Camellia.
REFERENCE	1 (sites)
AUTHORS	Kato, M., Mizuno, K., Crozier, A., Fujimura, T. and Ashihara, H.
TITLE	Caffeine synthase gene from tea leaves
JOURNAL	Nature 406 (6799), 956-957 (2000)
MEDLINE	20037335
REFERENCE	2 (bases 1 to 1438)
AUTHORS	Kato, M., Mizuno, K., Ashihara, H. and Fujimura, T.
TITLE	Direct Submission
JOURNAL	Submitted (16-AUG-1999) to the EMBL/GenBank and DDBJ databases Mizuno, University of Tsukuba, Institute of Agricultural and Forest Engineering, 1-1-1 Temoudai, Tsukuba, Ibaraki 305-8572, Japan (E-mail: koumou@sakura.cc.tsukuba.ac.jp, Tel:81-298-53-4656, Fax:81-298-55-2203)

FEATURES	SOURCE	Location/Qualifiers
phone		1. 1438
		/organism "Camellia sinensis"
		/db_xref "taxon:4442"
		90. 1149
		/name "TCS1"

QY	301	gaatgaatggaagaagaatggaaggaaatlaaataatgaagaagatggaagatctcaatttact	360
DB	331	GAATGATGGAAGAAATAAATGTAAGCAATTTAAATGGTAACATCTGAACTGCACTTTACT	390
QY	361	gaatgaatctctcttggagaatgaatgaagaactctctgaagacctatgctgaagtta	420
DB	391	TGAATGATCTCTTTTGGAAATGATTCATAGCTCTTTTAAAGGCTCTGCTGTGAGGTTA	450
QY	421	tttgtaacaagaatggaagaagttcagtgatagtgaaggaaataccagagctcttccatg	480
DB	451	TTTGTAATTAATTTGAGCAAGCTGCTGGTATGCTATGCTGCAACCGAGGCTCTTCATG	510
QY	481	gcgcgccttttccctgtaacagcttacaattgaatgaattctcttaccagttcatatgc	540
DB	511	GGCGGCTCTTTCTCTGTAAGAGATGATTAACATTAAGTCTCTCTCTCTCTCTCTCTCT	570
QY	541	ttactcagcagccaaagaagtaactcacaagcagcaggaagcttgcattaaacaaagaaga	600
DB	571	TTACTCGAGCAACCAAAAGCACTCTTAAGCAGAGAGGCTTGCTTTAAACCAAGCGCAAG	630
QY	601	tttctcgtct	660
DB	631	TTTCAATATCAAAACCAAAAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	690
QY	661	gaagatct	720
DB	691	AAGATTCGACAAATGTTCTCTCAATCTCAATGCGAAGAGGTGCTCTCTCTCTCTCT	750
QY	721	tttgtgaactctctgttgaagcaattctctcatctctcagaatgcaagtaagctgttactt	780
DB	751	TTTGTGATCTCTCTGTGATGCAATGTTCTGATCTCTTGAATATGCAAGGCTCTTACTT	810

QY 781 ggaactatagctatagccatgctgaacttgcttcaaggagttgatagatgaagata 840  
 |||||||  
 Db 811 GGGACACATATAGCATAGGCGATGCGAATGCGTTCACAGAGGATGATAGAGAGATA 870  
 |||||||  
 QY 841 aattagacccttcaatataccagctatttgatcacttaagaaatgaagataatag 900  
 |||||||  
 Db 871 AATTAGACACCTTCAATATATACAGCTATTATTCATATGAGGAAATGATATAG 930  
 |||||||  
 QY 901 tggagagagagagagatctctcaatctatctatctatctatctatctatctatctat 960  
 |||||||  
 Db 931 TGGAGAGAGAGAGAT 990  
 |||||||  
 QY 961 gaatgagagagagagatctctcaatctatctatctatctatctatctatctatctat 1020  
 |||||||  
 Db 991 AAT 1050  
 |||||||  
 QY 1021 tcaagagagagagagatctctcaatctatctatctatctatctatctatctatctat 1080  
 |||||||  
 Db 1051 TCAAGAGAGAGAGAT 1110  
 |||||||  
 QY 1081 aattgagagagagagatctctcaatctatctatctatctatctatctatctatctat 1140  
 |||||||  
 Db 1111 AATTGAGAGAGAGAT 1170  
 |||||||  
 QY 1141 tcaagagagagagagatctctcaatctatctatctatctatctatctatctatctat 1200  
 |||||||  
 Db 1171 TCAAGAGAGAGAGAT 1230  
 |||||||  
 QY 1201 ctatgagagagagagatctctcaatctatctatctatctatctatctatctatctat 1260  
 |||||||  
 Db 1231 CTATGAGAGAGAGAT 1290  
 |||||||  
 QY 1261 ggtcgaatataagagagatctctcaatctatctatctatctatctatctatctatctat 1320  
 |||||||  
 Db 1291 GGTCAAT 1350  
 |||||||  
 QY 1321 ctgagagagagagagatctctcaatctatctatctatctatctatctatctatctat 1380  
 |||||||  
 Db 1351 CTGAGAGAGAGAGAT 1410  
 |||||||  
 QY 1381 aaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaata 1408  
 |||||||  
 Db 1411 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1436  
 |||||||  
 RESULT 2  
 AB049752 1437 bp mRNA PLN 16-MAR-2001  
 LOCUS Atropa belladonna AbsAMT1.mrna.1  
 DEFINITION S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase,  
 complete cds  
 ACCESSION AB049752  
 VERSION AB049752.1 GI 13366160  
 KEYWORDS  
 SOURCE Atropa belladonna (strain M8) root cDNA to mRNA, clone lambda  
 ABSAMT1.  
 ORGANISM Atropa belladonna  
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; Solanales; Solanaceae; Atropa.  
 REFERENCE 1 (sites)  
 AUTHORS Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and  
 Yamakawa,T.  
 TITLE Cloning and expression of salicylic acid inducible and active  
 S-adenosyl-L-methionine:salicylic acid carboxyl methyltransferase  
 in transformed root culture of Atropa belladonna  
 JOURNAL: 2 (bases 1 to 1437)  
 REFERENCE Fukami,H., Asakura,T., Hirano,H., Abe,K., Shimomura,K. and  
 AUTHORS Yamakawa,T.  
 TITLE Direct Submission  
 JOURNAL: Submitted (10-OCT-2000) Takashi Yamakawa, the University of Tokyo,

Department of Global Agricultural Sciences: 1-1, Yayoi 1-chome,  
 Bunkyo-ku, Tokyo 113-8657, Japan  
 (E-mail:ayama@mail.ecc.u-tokyo.ac.jp, Tel:81-3-5841-7515,  
 Fax:81-3-5841-5304)  
 Location/Qualifiers  
 1..1437  
 /organism="Atropa belladonna"  
 /strain="M8"  
 /db\_xref="taxon:33113"  
 /clone="lambda AbsAMT1"  
 /tissue\_type="root"  
 52..1135  
 /gene="ABSAMT1"  
 62..1135  
 /gene="ABSAMT1"  
 /product="S-adenosyl-L-methionine:salicylic acid carboxyl  
 methyltransferase"  
 /protein\_id="BAB39396.1"  
 /db\_xref="GI:13366161"  
 /translation="MKVVEVLMHNGNDISYANNSIVQKVIIMTKPTEDQISDLY  
 CSFEPLFLIALCSSSANTLVELEVALEKRLHNSAGNLEHENDLPGNDP  
 NLFPSIKRFDQDPRQISEFGDFFSVSTPSTPIFVSFHFVHSSYIMLSQ  
 VPLLEKRNENIYIASISPEVYIKAYVQVKKDSNLEKYSFELMGKKVLELDR  
 ESDDPSKEQYIMELISMAINELVLELIEPKVDSNPIDQVPSPEVYIYERGS  
 SEINLEKALVPHKNSNGINDVNVAKCHPRAVPLLVNLEHQLMNVFYCYEEL  
 ISDLSKEKTEFTNVISLTKR"  
 ORIGIN  
 488 3 225 6 290 9 434 1  
 5907 Match 16.7% Score 498.2 PR 123 Length 1437  
 Best Local Similarity 54.9% Pred. No. 3,8641  
 Matches 599, Conservative 0, Mismatches 453, Indels 39, Gaps 5,  
 QY 80 gaatgagagagagatctctcaatctatctatctatctatctatctatctatctat 139  
 |||||||  
 Db 64 GAATGAGAGAGAT 123  
 |||||||  
 QY 140 ctctctctcaagagagatctctcaatctatctatctatctatctatctatctatctat 199  
 |||||||  
 Db 124 CTCTCTCTCAAGAGAGATCTCTCAATCTATCTATCTATCTATCTATCTATCTATCTAT 183  
 |||||||  
 QY 200 ggtcgaatataagagagatctctcaatctatctatctatctatctatctatctatctat 259  
 |||||||  
 Db 184 GGTCAAT 240  
 |||||||  
 QY 260 atcagagagagagagatctctcaatctatctatctatctatctatctatctatctat 319  
 |||||||  
 Db 241 ATCAAGAGAGAGAT 300  
 |||||||  
 QY 320 cagagagagagagatctctcaatctatctatctatctatctatctatctatctatctat 379  
 |||||||  
 Db 301 AAAAAAT 360  
 |||||||  
 QY 380 tgaatgagagagagatctctcaatctatctatctatctatctatctatctatctatctat 432  
 |||||||  
 Db 361 TGAATGAGAGAGAT 420  
 |||||||  
 QY 433 tgaatgagagagagatctctcaatctatctatctatctatctatctatctatctatctat 487  
 |||||||  
 Db 421 TGAATGAGAGAGAT 480  
 |||||||  
 QY 488 ttctccgagagagagatctctcaatctatctatctatctatctatctatctatctatctat 547  
 |||||||  
 Db 461 TTCTCCGAGAGAGAT 540  
 |||||||  
 QY 548 gtcagagagagagatctctcaatctatctatctatctatctatctatctatctatctat 607  
 |||||||  
 Db 541 GTCAAGAGAGAGAT 592  
 |||||||  
 QY 608 atcaagagagagagatctctcaatctatctatctatctatctatctatctatctatctat 667  
 |||||||





[illegible][illegible]





[illegible][illegible]





similarity but with ESI similarity are named as 'unknown' proteins. Genes without protein or EST similarity that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Alan Smith, <http://ftp.gene.washington.edu/PM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/MashU/ARI consortium for sequencing rat clones F6P23, F5B6, F17A6, and F1916, the ESSA group for sequencing clone F1304, and Scott Jackson, Jining Tian, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the Rice Bioinformatics Department, especially Lixin Zhou, Huiyi Kuang, Michael F. Brown, Eddy Lou, Tony Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: [atelig@r.ory](mailto:atelig@r.ory).

Location/Qualifiers

1. 93890

/organism="Arabidopsis thaliana"

/db\_xref="taxon 3702"

/chromosome="11"

/note="Sequence from clone F17024"

/repeat\_region

/gene="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/protein\_id="AAD19783.1"

/db\_xref="GI:4388828"

/translation="MNSNSVCAFPPEPPITGGPPVITFSSQTTITSSQCTAIIQD

/repeat\_region

/gene="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"

/note="At2g13900"













[illegible]

FEATURES	COMMENT
source	<p>Direct Submission</p> <p>Submitted (10-MAR-2000) MIPS at the Max-Planck-Institut fuer Biochemie, Am Klopferspitze 184, D-82152 Martinsried, FRG, E-mail: lemek@mips-biochem.mpg.de/mayer@mips-biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk</p> <p>Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at <a href="http://www.mips.biochem.mpg.de/proj/thal/">http://www.mips.biochem.mpg.de/proj/thal/</a> this fragment has an overlap with AtCHIVB84 at the 5' end and an overlap with AtRHIVB6 at the 3' end.</p> <p>Location/Qualifiers</p> <p>1. 198750</p> <p>/organism="Arabidopsis thaliana"</p> <p>/variety="Columbia"</p> <p>/db_xref="taxon:3702"</p> <p>/chromosome="4"</p> <p>5360 11095</p>
gene	<p>/gene="AT4g36190"</p> <p>3010(5400 5401 5500 5700 5801 6050 6121 6335, 6437 6553 6645 6800 6752 6838 6905 7171 7265 7359, 7418 7525 7653 7736 7854 7894 8706 8931 9012 9185, 9255 9450 9452 9458 9768 9823 9901 9972,10584 10657 10777 10865,10980 11095)</p>
cds	<p>/gene="AT4g36190"</p> <p>/note="similarity to F50f01.1, Caenorhabditis elegans, PAIX.G1688051</p> <p>contains ESTs:AL199435b.1, AA05515, H70809, AA50791, Z26781, AF965952.1, Z26780"</p> <p>/cdon_start=1</p> <p>/product="putative protein"</p> <p>/protein_id="CA805290.1"</p> <p>/db_xref="GI:7270572"</p> <p>/translation="MSSELGFALLFTFVLLSYLSFSSNGLLQPKRISHGLSKSNPCE SVELEINSLONSYQFLFLLEDPKRPQRYATYDILVPDPDIFLMIGEGEC NLTNNYISLAKKPKAGVATLEHYPGKSPKRYATKRLKYSSKALSPALAT YYQATCGVCLIMODSLNKNFENNSENVEWFFGYSQALSAMRLRLEPHITCSL ASSAVPVAVEPPEPDQIAESAPGPEETALDETINKLLEGLVNNPAKALNAEL DVADAEVLIADACVAMFIMQIDGCPNKLCPVIFACDNGDILNNPAKAVYRECM GVDSQSSKSTISPKRLILPAVILRSAPILMMPLVITVAYPLNARKISPSHJINRY HLIDCKSLPGKGVPEVDATILYSGDKILVSSHLILNFAIKLITGSDOPMHA SKSLSPILSPYIMTCHNIGHSLMGPISAAVIRGACNSSPAVKNPHEFEQF YVEYLDLRYPDGIEMKIGEGPCNGTIPNDYITVLAKEDAGVILEHRYKSSPF KSLATENIKYLSKVALHLIAEPQYIYASTLSMDSLNKFNPSGVENWFFPGASY SGATISAMPRLKRPPIITGSIASSAVPAVEPPEPDQIGESAPPRKALDETNL ELGIRKNNPAKALFNALIDVADELYLLAEAEVMAIYGNPLKCLVLEAEQRKD DLVPEPLSSOMHRIISVMHNRPHLNLNKEFLDITNLITGQVSTVIRCHNG HGSILMGPISAAVIRGACNSSPAVKNPHEFEQFIMSPGYSIRSSM"</p> <p>5360 5481</p>
exon	<p>/gene="AT4g36190"</p> <p>/number=1</p> <p>5482 5579</p>
intron	<p>/gene="AT4g36190"</p> <p>/number=1</p> <p>5580 5590</p>
exon	<p>/gene="AT4g36190"</p> <p>/number=2</p> <p>5791 5890</p>
intron	<p>/gene="AT4g36190"</p> <p>/number=2</p> <p>5891 6050</p>
exon	<p>/gene="AT4g36190"</p> <p>/number=3</p> <p>6051 6330</p>
intron	<p>/gene="AT4g36190"</p> <p>/number=3</p> <p>6331 6335</p>
exon	<p>/gene="AT4g36190"</p> <p>/number=4</p> <p>6336 6436</p>
intron	<p>/gene="AT4g36190"</p> <p>/number=4</p>



GenCore version 4.5  
Copyright (c) 1993 - 2000 GenCore Inc

OM protein - protein search, using sw model

Run on: August 31, 2001, 15:00:32 : Search time 59.12 Seconds

(without alignments)  
796,655 Million cell updates/sec

File: US-09-577-657A-1

Perfect score: 1847

Sequence: 1 FMNGBESSYAAQSSPTQ2

Scoring table:

Rossm62  
Gapop: 10 0 : Gapext 0.5

Searched: 425026 seqs, 132305027 residues

Total number of hits satisfying chosen parameters: 4,5026

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing:

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

1: SPTRMBL\_16:  
2: sp.archaea:  
3: sp.bacteria:  
4: sp.fungi:  
5: sp.human:  
6: sp.invertebrate:  
7: sp.mammal:  
8: sp.mhc:  
9: sp.oraneller:  
10: sp.phage:  
11: sp.plant:  
12: sp.podent:  
13: sp.povestit:  
14: sp.viruses:

Fred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No	Score	Query Match	Length	Fr	IT	Accession
1	1847	100.0	369	10	09FZNB	Q9FZNB
2	724	39.2	332	10	Q9SBK6	Q9SBK6
3	713	38.6	359	10	Q9SPV4	Q9SPV4
4	637.5	34.5	354	10	Q9F176	Q9F176
5	630	34.1	371	10	Q92234	Q92234
6	621	33.6	364	10	Q9FV79	Q9FV79
7	585.5	31.7	363	10	Q9LS20	Q9LS20
8	582.5	31.5	368	10	Q9JRL5	Q9JRL5
9	570	30.4	359	10	Q9ZPT3	Q9ZPT3
10	499.5	27.0	351	10	Q9X157	Q9X157
11	496.5	26.9	348	10	Q9FXC4	Q9FXC4
12	485.5	26.1	348	10	Q9FY76	Q9FY76
13	482.5	26.1	359	10	Q9LS10	Q9LS10
14	481	26.0	336	10	Q9F1N8	Q9F1N8
15	480	26.0	379	10	Q9FY73	Q9FY73
16	480	26.0	319	10	Q9F592	Q9F592
17	475.5	25.7	323	10	Q9FWP8	Q9FWP8
18	475.5	25.7	361	10	Q9FKR0	Q9FKR0
19	475.5	25.7	361	10	Q9FKR0	Q9FKR0

20	466.5	25.3	380	10	Q9FWP1	Q9FWP1
21	454	24.6	362	10	Q9FKR0	Q9FKR0
22	445.5	23.5	359	10	Q91MA2	Q91MA2
23	421	23.0	352	10	Q9M4F2	Q9M4F2
24	387	21.0	318	10	Q9FT13	Q9FT13
25	105.5	5.7	100	10	Q9LS11	Q9LS11
26	101.5	5.5	1016	3	Q9HT45	Q9HT45
27	99	5.4	319	5	Q9Z612	Q9Z612
28	98	5.3	1216	5	Q44875	Q44875
29	97	5.3	1216	5	Q44875	Q44875
30	96	5.2	1280	2	Q9ZAJ5	Q9ZAJ5
31	96	5.2	1682	10	Q9S8M5	Q9S8M5
32	94.5	5.1	469	14	Q9W852	Q9W852
33	94.5	5.1	469	14	Q9W852	Q9W852
34	94.5	5.1	787	1	Q58025	Q58025
35	94.5	5.1	725	14	Q9F179	Q9F179
36	94.5	5.1	442	14	Q9W854	Q9W854
37	94	5.1	983	10	Q9ZUK3	Q9ZUK3
38	93.5	5.1	525	3	Q9Z830	Q9Z830
39	93	5.0	245	2	Q9XMK6	Q9XMK6
40	93	5.0	735	5	Q25855	Q25855
41	93	5.0	749	5	Q9Z602	Q9Z602
42	93	5.0	847	10	Q91H00	Q91H00
43	92.5	5.0	2185	14	Q9E7C3	Q9E7C3
44	92.5	5.0	2185	14	Q9E7C3	Q9E7C3
45	92	5.0	392	3	Q94240	Q94240

#### ALIGNMENTS

RESULT 1  
Q9FZNB 1 PPIIMINAVY: PPT: 369 AA.

AC Q9FZNB: PPIIMINAVY: PPT: 369 AA.  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE CAFFEINE SYNTHASE.  
GN TCS1.  
OS Camellia sinensis (Tea).  
OC Eurythya, Viridiplantae, Eurythya, Eurythya, Eurythya.  
OC Magnoliophyta, Eudicotyledons, Core eudicotyledons, Asterales, Ericales.  
OC Theaceae, Camellia.  
OX NCBI\_TaxID=4442.  
PN [1]  
PP SEQUENCE FROM N.A.  
RA MEDLINE=20437335; PubMed=10984041;  
RA Kato M., Mizuno K., Crozier A., Fujimura T., Ashihara H.:  
RT "Caffeine synthase gene from tea leaves."  
RU Nature 406:956-957(2000).  
DR EMBL: AB031280; EMBL2278.1; -  
SQ SEQUENCE: 369 AA: 4127 MW: 41055.8874/5576.1664:

Query Match 100.0% Score 1847: OR 19: Length 369:  
Best local similarity: 100.0%: Pred. No. 1, 1e-144:  
Matches 455: Conservative 0: Mismatches 0: Gaps 0:  
Q7 1 FMNGBESSYAAQSSPTQ2VASMADPALENAVETLPSMDPHLOAINADLCAAGPNTF 60  
|||||  
Q7 14 FMNGBESSYAAQSSPTQ2VASMADPALENAVETLPSMDPHLOAINADLCAAGPNTF 73  
|||||  
Q5 1 AVTSTFEMKPKFRETNETFLQVYNTLEFNGRRIHFKLSSVVGSRLEFVGVVM 120  
|||||  
Q7 74 AVTSTFEMKPKFRETNETFLQVYNTLEFNGRRIHFKLSSVVGSRLEFVGVVM 133  
|||||  
Q7 121 GVPSTFEMKPKFRETNETFLQVYNTLEFNGRRIHFKLSSVVGSRLEFVGVVM 180  
|||||  
Q7 134 GVPSTFEMKPKFRETNETFLQVYNTLEFNGRRIHFKLSSVVGSRLEFVGVVM 193  
|||||  
Q7 181 AVTSTFEMKPKFRETNETFLQVYNTLEFNGRRIHFKLSSVVGSRLEFVGVVM 240  
|||||

```

104 AVLSQPHLDEFTMFLNAPSQVFNPCVWVLLPZCSPDSMOSCTFWELLAMATAELVS 253
105
106 241 OGILDEKIDITFNIPSPASLEEVKQIVEDDSPTIDIRFTDINSEVNGENKVVPGPK 400
107
108 254 OGILDEKIDITFNIPSPASLEEVKQIVEDDSPTIDIRFTDINSEVNGENKVVPGPK 413
109
110 401 FTKVVAATPELISNORGEIMKIKRFTHIVASLEKLEKTSIIIVLSKIDG 356
111
112 414 FTKVVAATPELISNORGEIMKIKRFTHIVASLEKLEKTSIIIVLSKIDG 369
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500

```

```

501 MAY-2000 (Trembl, 13, last sequence update)
502 01-MAY-2000 (Trembl, 13, last annotation update)
503 S-ADENOSYL-L-METHIONINE:GLYCILIC ACID CARBOXYL METHYLTRANSFERASE
504 SAMT.
505 Clarks brewer.
506 Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:
507 Magnoliophyta: eudicotyledons, core eudicots, Rosidae, eurosids 11:
508 Myrtales: Onagraceae: Clarkia.
509 NCBI_TaxID=36903;
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```



HA Tabata S.:  
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. VI.  
 RT Sequence features of the regions of 1,367,183 bp covered by 14  
 RT physically assayed P1 and PAC clones."  
 DR DNA Res. 5:203-216(1998).  
 DR EMBL: A013449; FAF10919.1; -

SO SEQUENCE 354 AA: 40033 MW: 9519636DCACAD329 CRC64;

Query Match 34.58, Score 637.5, DB 10, Length 354;  
 Best Local Similarity 39.7%, Pred. No. 9, 7, 4;  
 Matches 145; Conservative 64, Mismatches 131, Indels 25, Gaps 9;

QY 2 MNRGSSVAQNSSTFOQVASMAGPAPL-ENAVETI-PSRPHHQAIMNADLGCAGAPNTE 60  
 DB 1 MSGRQGNSTNSITLQKVLKAPKPYLVNTGMMINLNP-NIKVADLGCATCENEF 59  
 QY 61 AVISTIKMMKRCRELNCQLELVYVNDLPNDENTLFRGL---SSEVIGKCEVPC 117  
 DB 60 LTMALVNTLVNLCQCNKRPETDCLNLPNDENTLFRKVPFKKVKSPKPL- 115  
 QY 118 YVMGVPSSPFRPLFPNSLHVHSSVYVHMLTGAPKGLTKEGLANKKGIYSKTSPPV 177  
 DB 116 FVSGVPSSPSPRIFFKSLHFVHSSVSLHMLSKVPGLEK-----NSSVYITTSPPN 169  
 QY 178 VREAVLSQFHEPFTPLNASSQFVPCVWVLLPGPG-CSDPSDMSQCFETWELAMATA 236  
 DB 170 AYKAVLNQFQSDKSEFLKMSFEMVSNQPVLLFTFGKTLDPGLHDCHEWLLSTSLR 229  
 QY 237 ELVSQGLIDELKLDFTNIPSYFASLEVKDIVERDGFYIDHLE--GFDL-----DSVHM 289  
 DB 230 DLVEGGLVSAKSDVSNIPEDYDPSKEVEMDINEGSEFINDLEHGFELGSLNHDEQYM 289  
 QY 290 GFNKKWPCKEKFTVVAFAFPIISNQFPEIMKIKYFPTTIVVSDTAKLKTPTTIT 349  
 DB 290 LHSQISAGQFANCTRAVSHSMVALFQVDMLETKRFAHV--SGHASTKTKTIVII 347  
 QY 350 VLKRI 354  
 DB 348 VVSLI 352  
 RESULT 5  
 ID 023234 PRELIMINARY; PPT: 371 AA.  
 AC 023234;  
 DT 01-JAN-1998 (TEMBLrel. 05, Created)  
 DT 01-MAR-1999 (TEMBLrel. 10, Last sequence update)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
 DB HYPOTHEICAL 42.0 KDA PROTEIN.  
 DB C7A10.890 OR AT4G36470  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:  
 OC Magnoliophyta: Euphorbiales: Caryophyllales: Rosales: Violaceae: Vi-  
 OC brassicales: Brassicaceae: Arabidopsis.  
 OC NCBI\_TaxID=3702;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Bryan M. Terryn N. Vos P. Huijgen L. Meuwis H.W. Scholten C.  
 RA Chabotz N.;  
 PI Submitted (Dec-1998) to the EMBL/GenBank/DBI databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR 2000) to the EMBL/GenBank/DBI databases.  
 DR EMBL: Z99708; CAB16845.1; -  
 DR EMBL: A1161589; CAB80313.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 371 AA: 41984 MW: F5PR341EASQ6A9D CPO64.

Query Match 34.18; Score 630; DB 10; Length 371;  
 Best Local Similarity 37.38; Pred. No. 4, 3e-44;

Matches 135; Conservative 84; Mismatches 131; Indels 12; Gaps 5;

QY 1 FMRGSSVAQNSSTFOQVASMAGPAPL-ENAVETI-PSRPHHQAIMNADLGCAGAPNTE 60  
 DB 1 YMGDGRKISTAPNSITQKRASTAKHITFTLQGLY-KEIPFSLGIMLQSSSPNTE 69  
 QY 61 AVISTIKMMKRCRELNCQLELVYVNDLPNDENTLFRGL---SSEVIGKCEVPC 117  
 DB 70 STTDFTRVAVAHNRPILPPLPFPSTPLNLPNDENTLFRKVPFKKVKSPKPL- 129  
 QY 112 GFVPCVVMGVPSSPFRPLFPNSLHVHSSVYVHMLTGAPKGLTKEGLANKKGIYS 171  
 DB 130 GPVY--FLAAPGSPGVPGLPPEHTHFVYASSTHMLSNVPLATVYQDGSINQCVSLC 187  
 QY 172 KTSFVVFVAVTSQFHEPFTPLNASSQFVPCVWVLLPGPG-CSDPSDMSQCFETWEL 231  
 DB 188 SLSEAVSKAVSQFKEPDESTFLPCKSEKVSAGPVLLITIGEGPPHYVPCNSFEWEL 247  
 QY 232 AMALVLSQGLIDELKLDFTNIPSYFASLEVKDIVERDGFYIDHLE--GFDL-----DSVHM 289  
 DB 248 SRSLADLVAGQFTEFEKIDSYDMHFPAPRADTEGYNKSGSPFLFRLMLEVKKDKNT 307  
 QY 292 NKKWPCKEKFTVVAFAFPIISNQFPEIMKIKYFPTTIVVSDTAKLKTPTTIT 351  
 DB 308 FQITSYKAVAKIVFAVQSSMLVQHPRKILKLFHYCPMWDELAKDILPILVVL 367  
 QY 352 SK 353  
 DB 368 RK 369  
 RESULT 6  
 ID 09FY29 PRELIMINARY; PPT: 364 AA.  
 AC 09FY29;  
 DT 01-MAR-2001 (TEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TEMBLrel. 16, Last annotation update)  
 DB SAM-BENZOIC ACID CARBOXYL METHYLTRANSFERASE.  
 DB CN  
 OS Antirrhinum majus (Garden snapdragon).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:  
 OC Magnoliophyta: Eudicotyledons: Core eudicots: Asteridae: euasterids 1:  
 OC Lamiales: Scrophulariaceae: Antirrhinum.  
 OC NCBI\_TaxID=4151;  
 RN 111  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20312758; PubMed=10852939;  
 RA Dubrovic N., Morfitt J.M., Martin C.J., Gledhill N., Kolesova N.,  
 RA Kish C.M., Bonham C., Wood K.;  
 RT "Developmental regulation of methyl benzoate biosynthesis and emission  
 RT in Snapdragon flowers."  
 EL Plant Cell 12:949-961(2000).  
 LE EMBL: A136452; AA:05284.1; -  
 KW Transferase, Methyltransferase.  
 SO SEQUENCE 364 AA: 41010 MW: C10C8E8D4A581419 CRC64;

Query Match 34.68; Score 621.5, DB 10; Length 364;  
 Best Local Similarity 38.18; Pred. No. 2, 1e-43;  
 Matches 140; Conservative 64, Mismatches 128, Indels 35, Gaps 8;

QY 5 GERSSVAQNSSTFOQVASMAGPAPL-ENAVETI-PSRPHHQAIMNADLGCAGAPN 58  
 DB 14 GTSSTVANSNSLQKVMKSLHVLPIIKLIIQHVHVPKPKKMM-----LMSGSSSPN 68  
 QY 59 TFAVSTIKMMKRCRELNCQLELVYVNDLPNDENTLFRGLSSEVIGKCEVPC 117  
 DB 69 ALLVMSCLINTEIDLYTEKINLEFEVHNLNDPNDNRLIKLSH RGN ---C 121  
 QY 118 YVMGVPSSPFRPLFPNSLHVHSSVYVHMLTGAPKGLTKEGLANKKGIYSKTSPPV 177  
 DB 122 FVGLPSSPFGILLPKKSLHFVHSSVSLHMLSKVPGLEK-----NNOIYMAIESPPE 176

178 VREAVISQPHHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWELAMAAAE 237  
 177 VYKAAVAYKEDPSTFKLKEEIVPGKAVITFNSEVSDPSKDIATITLAKTIVE 236  
 238 IVSQGIIDKRPENIPSEFASLEEFKQDVEPDSFTLIDHECHPI--DSVHMOEND-- 293  
 237 MWAGGLAKMDLSPNPISPIPSCTPVEFAALISFNSFTILKLVFVWVWASVYTLQDQ 296  
 294 -----KWKGEKTKVAPATPEPTISNQGPEIMDKYDKPTHVSDLEAKLPKTT 346  
 297 QDPSTFKQCSKTFVALCVPAITEPMASHFTSTIMETFGYAKKIVFHSVENSSTVS 356  
 347 IIVLSK 354  
 357 IIVLSK 363

## RESULT 7

Q9LS20

ID Q9LS20 PRELIMINARY: PRT: 363 AA.

AC Q9LS20: 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE-  
 LIKE.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae; eurosids II;  
 OC Brassicaceae; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID:3702.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA:  
 RA Kaneko T., Katoh T., Asamizu E., Sato S., Nakamura Y., Kotani H.,  
 Tabata S.:  
 KI "Structural analysis of Arabidopsis thaliana chromosome 5. XI."  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB028606; BAA97534.1;  
 DR InterPro: IPR000794;  
 DR PROSITE: PS00606; L\_KELUACYL\_SYNTHASE; UNKNOWN.  
 SO SQUENGE 363 AA; 41154 MW; 406161084B34108 CPG64;

Query Match: 31.7%; Score 585.5; DR 10; Length 363;  
 Best Local Similarity: 35.7%; Pred. No. 2e-40;  
 Matches 128; Conservative 67; Mismatches 115; Indels 49; Gaps 9;

1 2 MNRGDSSTAAQNSFTQGVASMAQALNAVETLSPFHQALNAALDCAAGNFTA 62  
 1 MKGDSSTAAQNSFTQGVASMAQALNAVETLSPFHQALNAALDCAAGNFTA 60  
 62 VTS -IKMMKKKPFELNQTLEQVYINLFGNENFT-----KLSSTV 107  
 61 VMSLIVNTITTSQOKKKNPEIN-----CTNLPNNPNTTTPKIVPAHKLKMOV 113  
 108 IGKCFPVVYVWVPSFGHLPFPNSLHVSSSVHMLTQARKGLTSREGALNKCK 167  
 114 KGR-----CFISGVPSFSLPFSKSLHFNHSSLCIHLMSKVPGLID-----NKKN 161  
 168 IYISKTSPPVVPVAVISQPHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWEL 226  
 162 VYKAAVAYKEDPSTFKLKEEIVPGKAVITFNSEVSDPSKDIATITLAKTIVE 220  
 227 TWELAMAAAEVSDPSFTFKLKEEIVPGKAVITFNSEVSDPSKDIATITLAKTIVE 280  
 223 MWAGGLAKMDLSPNPISPIPSCTPVEFAALISFNSFTILKLVFVWVWASVYTLQDQ 283  
 281 GDULOSVEQEND-----KWKGEKTKVAPATPEPTISNQGPEIMDKYDKPTH 341  
 281 FSVKTFTEVAKIDNINISQCPVFIKPASTIPSTITETMLAPHPDAIMVPEPYTY 349

## RESULT 8

Q9LPT3

ID Q9LPT3 PRELIMINARY: PRT: 368 AA.

AC Q9LPT3: 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
 DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE-LIKE  
 DE PROTEIN, FLUORAL NECTAR-SPECIFIC PROTEIN-LIKE  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;  
 OC Magnoliophyta; eudicotyledons; core eudicotyledons; Rosidae; eurosids II;  
 OC Brassicaceae; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID:3702.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA:  
 RA Sato S., Nakamura Y., Kaneko T., Katoh T., Asamizu E., Tabata S.:  
 KI Submitted (JUN 1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-COLUMBIA:  
 RA MEDLINE 26277486, PubMed 10819329;  
 RA Nakamura Y.:  
 KI "Structural analysis of Arabidopsis thaliana chromosome 3. I. Sequence  
 features of the regions of 4,504,864 bp covered by sixty p1 and TAC  
 clones."  
 RL DNA Res. 7:131-135(2000).  
 DR EMBL: AB028622; BAA91375.1;  
 DR InterPro: IPR000794;  
 DR PROSITE: PS00606; L\_KELUACYL\_SYNTHASE; UNKNOWN.  
 SO SQUENGE 408 AA; 42010 MW; 11405644144e CPG64;

Query Match: 31.5%; Score 582.5; DR 10; Length 368;  
 Best Local Similarity: 37.7%; Pred. No. 3.6e-40;  
 Matches 133; Conservative 65; Mismatches 118; Indels 37; Gaps 9;

1 2 MNRGDSSTAAQNSFTQGVASMAQALNAVETLSPFHQALNAALDCAAGNFTA 60  
 1 MKGDSSTAAQNSFTQGVASMAQALNAVETLSPFHQALNAALDCAAGNFTA 64  
 61 AVSTIKMMKKKPFELNQTLEQVYINLFGNENFT-----GLSSHYGNKK 112  
 65 LVMSEIVNTITTSQOKKKNPEIN-----CTNLPNNPNTTTPKIVPAHKLKMOV 121  
 113 EGVCTPVVYVWVPSFGHLPFPNSLHVSSSVHMLTQARKGLTSREGALNKCK 172  
 122 IYISKTSPPVVPVAVISQPHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWEL 171  
 174 IYISKTSPPVVPVAVISQPHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWEL 232  
 172 IYISKTSPPVVPVAVISQPHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWEL 231  
 233 MAIAIVSQGIIDKRPENIPSEFASLEEFKQDVEPDSFTLIDHECHPI--DSVHMOEND-- 285  
 232 IYISKTSPPVVPVAVISQPHPTMTNANSPVVPNGCVLLLPKQCSPTDMSCTTWEL 291  
 286 SVENGENKMKVREKPKTKV-----KATFTHLSNQGPEIMDKYDKPTH 331  
 292 HSFVAKIDNINISQCPVFIKPASTIPSTITETMLAPHPDAIMVPEPYTY 344  
 RESULT 9  
 Q9LPT3  
 ID Q9LPT3 PRELIMINARY: PRT: 359 AA.  
 AC Q9LPT3:  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE P8822.17.

OS Arabidopsis thaliana (Mouse tail cross).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta  
 FA Magnoliophyta: euembryales: core eudicots: Rosidae: eurosids II  
 OC Brassicales: Brassicaceae, Arabidopsis.  
 RX NCBI\_TaxID=3702;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., Vaananen S.E.,  
 Barstead M.E., Mason T.M., Bowman C.L., Pechen G.M., Bhatia M.,  
 Carreira A.T., Crosby T.H., Fawcett G.P., Tward D., Nierman W.C.,  
 Fraser G.M., Venter L.C.,  
 RT "Arabidopsis thaliana chromosome II BAC F9822 genomic sequence."  
 RI Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases  
 DR EMBL: AC006528; AAD19741.1;  
 DR InterPro: IPR000794;  
 DR Prosite: PS00606; K:RELACYL\_SYNTHASE; UNKNOWN\_1.  
 SQ SEQUENCE 359 AA: 41370 MW: 44940.94 kDa: PHEHLDQ654;

Query Match 30.9% Score 570; DB 10; Length 359;

Best Local Similarity 36.9%; Pred. No. 3 7e-39;  
 Matches 136; Conservative 64; Mismatches 147; Indels 32; Gaps 10.

DB 2 MNGGSPSSVAGNSSTFGVASMAGP-ALFNAVFETPSDFHIALNADYCGAGPTF 60  
 1 MKGTRHSHATNSHYRQSVFETLPLVTEVEMLNKGF-CTKVADPSTGQNTV 59  
 QY 61 AVISTIKPMKKKPELNCLELVYLNLEPGDNFLEK-----GLSEVIGNKE 113  
 DB 60 LAMSAIAYITIMSYGQMSKNPPIYNTIPENDITEFKLFHSFQKIKFVYKGR--- 116  
 QY 114 EVCQYMGVGVGSHGLFPNSHLVHSSVVMITQAKGLTSEGLAIKNGKIYISK 173  
 DB 117 ---WFSVGPSSVYSKLPKSLHVFHSAFSTHMSIPIDGLES-----NKSITHYV 167  
 QY 174 SPVVFAYLSGHEFTFTNAKSAQVVPNAGVYLPGQSPSSQSGP-TWELIA 232  
 DB 168 YPENVKSYTNCEKLIIFSLFKRSEEVYHNIMVITPGEKVSITLS-KLEHYWLSLS 226  
 QY 233 MAIAELVSGQLIFKGLTFNIPSYFASLEEVKQIVEPDSTIDHIEGFD LGS 286  
 DB 227 DCLLDLASEGVNDSVYKSNMFPYFNNEEYVEFLKSGSEFIKIKFIDHVPYKIDR 286  
 QY 287 VEQEEND--KWRGKFTVVFATPEILISNQPPEIMKIDKTHIVSULE-AKLPR 343  
 DB 287 EEDDEBSIDLEAGLKAASWAKITPELVANFGAILEVFNKYAHYAKYLSVNHPR 346  
 QY 344 YTSILEVLS 352  
 DB 347 NMTLIVVS 355

RESULT 10

DB 09X157 PRELIMINARY: PRT: 351 AA  
 AC 09X157;  
 DT 01-NOV-1999 (TREMBLrel 12, Created)  
 DT 01-NOV-1999 (TREMBLrel 12, last sequence update)  
 DT 01-NOV-1999 (TREMBLrel 12, last annotation update)  
 DE P9L1.6  
 GN P9L1.6  
 OS Arabidopsis thaliana (Mouse tail cross).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta  
 FA Magnoliophyta: euembryales: core eudicots: Rosidae: eurosids II;  
 OC Brassicales: Brassicaceae, Arabidopsis.  
 RX NCBI\_TaxID=3702;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC STRAIN=CV, COLUMBIA;  
 RA Vrsoskaja V.S., Schwartz I.F., Yu G., Toriumi M., Tong G., Liu S.,  
 Liu J., Liu A., Li J., Krumholzskaja L., Juras J., Gonzalez A.,  
 Altati H., Araujo R., Brooks S., Buehler E., Chao G., Conn L.,

RA Gouway A.B., Dunn P., Hansen N., Hutzot L., Khan S., Kim C., Palm C.,  
 PA Fowler D., Shien P., Walker M., Norris P.W., Parker J.R.,  
 FA Fud-Rupel N.A., Theologis A;  
 RT "Arabidopsis thaliana chromosome I BAC P9L1 sequence."  
 RI Submitted (JUN 1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC007591; AAD39641.1;  
 SQ SEQUENCE 351 AA: 38851 MW: 45444.7326578AFL CHG64;

Query Match 27.0% Score 439.5; DB 10; Length 351;

Best Local Similarity 33.6%; Pred. No. 3 4e-33;  
 Matches 122; Conservative 73; Mismatches 141; Indels 27; Gaps 10;

DB 2 MNGGSPSSVAGNSSTFGVASMAGP-ALFNAVFETPSDFHIALNADYCGAGPTF 59  
 1 MKGTRHSHATNSHYRQSVFETLPLVTEVEMLNKGF-CTKVADPSTGQNTV 60  
 QY 60 FAVISTIKPMKKKPELNCLELVYLNLEPGDNFLEKGLSLSEVIGNKEEVPDY 118  
 DB 61 FAVDTITQALYKFTSTIPNTTTPQGVFNVSSTPNALEFALP-----PQPPYF 112  
 QY 119 VMCVPSFPHGLFPNSHLVHSSVVMITQAPKGLTSEGLAIKNGKIYISKTPPV 178  
 DB 113 VAGVPSFPGNI FPKAHINIAVSSCALCMI SPTPTTSSPAVNPCHYTHASAFVA 172  
 QY 179 EPAYLSGHEFTFTNAKSAQVVPNAGVYLPGQSPSSQSGP-TWELIAM 233  
 DB 173 -VAVSSQYKRDIKLFHASSQELAEKGMALLVPGVPGDFLCQFASIGSE--FDLGS 228  
 QY 234 AVAFVSGQLIFKGLTFNIPSYFASLEEVKQIVEPDSTIDHIEGFDSEVQEPND 293  
 DB 229 GIDMAKKEGIIIEEVNSNLPITYTTKELEDIIRSNKELDKME--TIGSDAD--D 284  
 QY 294 KWRGKFTVVFATPEILISNQPPEIMKIDKTHIVSULE--AFKIPKTTITIV 350  
 DB 286 TMDHPSFPIYI FAVLEFIVRTHFGHIIIDLFYVAKIKASSPTIQQPQTHIMFAL 344  
 QY 351 LSR 353  
 DB 345 LSR 347

RESULT 11

DB 09FYCA PRELIMINARY: PRT: 348 AA  
 AC 09FYCA;  
 DT 01-MAR-2001 (TREMBLrel 16, Created)  
 DT 01-MAR-2001 (TREMBLrel 16, last sequence update)  
 DT 01-MAR-2001 (TREMBLrel 16, last annotation update)  
 DE ATP-LIKE PROTEIN.  
 GN P28D10.50.  
 OS Arabidopsis thaliana (Mouse tail cross).  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta  
 FA Magnoliophyta: euembryales: core eudicots: Rosidae: eurosids II;  
 OC Brassicales: Brassicaceae, Arabidopsis.  
 RX NCBI\_TaxID=3702;  
 RN [1]  
 RP SOURCE FROM N.A.  
 RC Tolsen M., Berger C., Cooke P., Grellet F., Laurie M., Meves H.W.,  
 PA Boud S., Lemcke F., Mayor K.F.X., Querter P., Salazar M.;  
 RI Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: A591254; CAC03346.1;  
 SQ SEQUENCE 348 AA: 38515 MW: 30099.982603E77 CHG64;

Query Match 26.9% Score 426.5; DB 10; Length 348;  
 Best Local Similarity 34.4%; Pred. No. 4 3e-33;  
 Matches 121; Conservative 67; Mismatches 121; Indels 43; Gaps 10;

```

UY 5 GPCSSAONSSFTQGVASMAQALFNAVEITLPSRPHILALNALALDGAAPNFFAVIS 64
   111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 10 GRPNSTYHFKSKYQVAVIAAKERINAEALSTKLDLDFSLNINADFGSSQVNFIAVQ 69
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
UY 65 TIKRMMERKQ-PRI NQUTLELVAYNLNLENDENTL PKRI SSEVITONKFFVPCVVMGP 123
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 70 TLIDAVNKYKRESNIEGLFQVGFNNSSNDENTLTKLPARL-----YASGVIP 121
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 124 GSFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 122 GSFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 184 SDFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 180 GYKILVGSFLTARAGELVSGILLILLSGR-----PFGVMEIYEGMIDPILSSLMEL 235
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 239 VSGITLTPKTLFENLDPSPASTFVYKIVPPGQSTID-----HIEG-EDLISVEKCE 291
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 236 ANGLIDOCKIIPKILFVAPVDFLNLJLELNKCLLEAFERKISHAGCEYLLDF----- 290
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 292 NDKWVEKREKTKVAVAFTEFLISNCFPEIMKRYKFIHIVASLFAATIPK 343
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 291 -----EYLSAKVIVGSSVSLFGUGMEKTE-----LVKERTQEMLPQ 331
   111 111 111 111 111 111 111 111 111 111 111 111 111

```

## RESULT 12

UYFC6 PRELIMINARY: PRT: 348 AA.

```

ID 09FC6:
AC 09FC6:
DT 01-MAR-2001 (TREMURel. 16, Created)
DT 01-MAR-2001 (TREMURel. 16, last sequence update)
DT 01-MAR-2001 (TREMURel. 16, last annotation update)
DE PROTEININASE ATP-LIKE PROTEIN
CN F28D10.30.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicotids; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RA Delaney M., Berger G., Cooke R., Grellet F., Laurie M., Mendes H.W.,
RA Rodd S., Lemcke K., Mayer K.F.X., Quetzel F., Salanoubat M.;
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.
RN 121
RP SEQUENCE FROM N.A.
RA Arabidopsis sequencing project;
RA Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases
DR EMBL; AL391254; CAC01534.1; -.
KW Kinase.
SQ
SEQUENCE 448 AA: 48940 MW, 4160CFPEVETAT4 TCT4.

```

Query Match 26.1%; Score 485.5; DB 10; Length 348,  
Host Local Similarity 44.9%; Pred. No. 3; Size 32,  
Matches 128; Conservative 65; Mismatches 129; Indels 45; Gaps 13;

```

UY 5 GPCSSAONSSFTQGVASMAQALFNAVEITLPSRPHILALNALALDGAAPNFFAVIS 64
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 10 GRPNSTYHFKSKYQVAVIAAKERINAEALSTKLDLDFSLNINADFGSSQVNFIAVQ 69
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 65 TIKRMMERKQ-PRI NQUTLELVAYNLNLENDENTL PKRI SSEVITONKFFVPCVVMGP 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 70 TLIDAVNKYKRESNIEGLFQVGFNNSSNDENTLTKLPARL-----YASGVIP 121
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 124 GSFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 122 GSFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 184 SDFHGLFPPNSIIVHSSSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 123
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 180 GYKILVGSFLTARAGELVSGILLILLSGR-----PFGVMEIYEGMIDPILSSLMEL 235
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 239 VSGITLTPKTLFENLDPSPASTFVYKIVPPGQSTID-----HIEG-EDLISVEKCE 291
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 236 ANGLIDOCKIIPKILFVAPVDFLNLJLELNKCLLEAFERKISHAGCEYLLDF----- 290
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 292 NDKWVEKREKTKVAVAFTEFLISNCFPEIMKRYKFIHIVASLFAATIPK 343
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 291 -----EYLSAKVIVGSSVSLFGUGMEKTE-----LVKERTQEMLPQ 331
   111 111 111 111 111 111 111 111 111 111 111 111 111

```

```

UY 245 IALVLSUGLIDRUKDFTNIPSYFASLEPVKDIVRDSFTLIDHIEGDLSSVMKGRHK 294
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 232 LNDIAGGLIDRUKDFTNIPSYFASLEPVKDIVRDSFTLIDHIEGDLSSVMKGRHK 285
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 295 WVGFEFTFY-VPAPEFTISMVGFPEIMKRYKFIHIVASLFAATIPK 343
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 286 IPIIDPEFTIYSEFVTVGGIVASFGHGVMEKTE-----VVKIKTEPMI.POLANKPCMOY 341
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 246 IALVLSI 354
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 342 LIVLEPI 348
   111 111 111 111 111 111 111 111 111 111 111 111 111

```

## RESULT 13

UYLS10 PRELIMINARY: PRT: 459 AA.

```

ID 09LS10:
AC 09LS10:
DT 01-OCT-2000 (TREMURel. 15, Created)
DT 01-OCT-2000 (TREMURel. 15, last sequence update)
DT 01-MAR-2001 (TREMURel. 16, last annotation update)
DE S-ADENOSYL-L-METHIONINE:SALICYLIC ACID CAGEXYL METHYLENE-PPINASE-
DE LIKE
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; core eudicotids; Rosidae; eurosids II;
OC Brassicales; Brassicaceae; Arabidopsids.
CX NCBI_TaxID=3702;
RN 111
RP SEQUENCE FROM N.A.
RA STRAIN=COLUMBIA;
RA Kaneko I., Katoh I., Asamizu H., Sato S., Nakamura Y., Kotani H.,
RA Tabata S.;
RL Structural analysis of Arabidopsis thaliana chromosome 5. XI.*;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028006; BA097544.1; -.
SQ
SEQUENCE 553 AA: 40344 MW, 456527A0950D34 CROG4;

```

Query Match 26.1%; Score 482.5; DB 10; Length 459,  
Host Local Similarity 33.7%; Pred. No. 6; Size 32,  
Matches 115; Conservative 65; Mismatches 124; Indels 47; Gaps 11;

```

UY 2 MNGGESSAONSSFTQGVASMAQALFNAVEITLPSRPHILALNALALDGAAPNFFAVIS 56
   111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB 10 MNGGESSAONSSFTQGVASMAQALFNAVEITLPSRPHILALNALALDGAAPNFFAVIS 56
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 57 PNF-----AVISTIKPMKPKPIINQCTIPIGVYINLPINPNTI PK--GLSSEVIGN 110
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 67 PNFVHAGSTIDIVAKSKRLESTENSIVPLFQVGFNNQINDPNTI PKRQYLS----- 120
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 112 KLELVYVMAVISEHDLFPNLSLHVSYSVHMLTQAPKQLTSRHLANCKYIVTSKTSPPVPAVL 170
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 121 --PEPEFVSQVGSFVGLPPLNSIHGHTSYTHMLSKYVQVWIKRSMANNKYNLGC 178
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 171 SKRTIVVPEFVYI SQFHDPFMR NARSGIVVPGQVAVILFPGQGVVSGVGTWIC 230
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 179 NNLLEEVTK-AKVQVFIKMEIFLLAFAEELVPGGLIMVI---GEGL-PCGVSLYETWGG 234
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 231 IAW-----AIALVSGGLIDRUKDFTNIPSYFASLEPVKDIVRDSFTLIDHIEGDLSSVMKGRHK 294
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 234 YVMDTIGDLMMAKSGILSEKILFSLPVYTHGFSNFKERIKNSDFTLEME----- 288
   111 111 111 111 111 111 111 111 111 111 111 111 111
UY 296 SVPMENDRWVGEKTKVAVAFTEFLISNCFPEIMKRYKFIHIVASLFAATIPK 343
   111 111 111 111 111 111 111 111 111 111 111 111 111
DB 289 ITSHPLEGRPLINDFTITFPAFLITLIEKHGQDVVELF 329
   111 111 111 111 111 111 111 111 111 111 111 111 111

```

## RESULT 14

UYFC8 PRELIMINARY: PRT: 374 AA.

```

ID 09FC8:
AC 09FC8:
DT 01-MAR-2001 (TREMURel. 16, Created)
DT 01-MAR-2001 (TREMURel. 16, last sequence update)

```

D1 01-MAY-2001 (TREMUR: 16, last annotation update)  
 DE 3'-AEMOSYL-L-METHIONINE:SALICYLIC ACID CARBOXYL METHYLTRANSFERASE  
 DE PROTEIN.  
 OS Arabidopsis thaliana (Mustard cross)  
 OC Eukaryota: Viridiplantae: Embryophyta: Tracheophyta: Spermatophyta:  
 CC Magnoliopsida: Eudicotyledones: Core eudicotyledons: Rosidae: eucosids II:  
 OC Brassicales: Brassicaceae: Arabidopsis  
 CX NP\_174191.2:702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 FC STAIN-COLIMBIA  
 RX MEDLINE:98403884; PubMed-9734815;  
 PA Kevani H., Nakamura Y., Sato S., Asamizu F., Kaneko T., Miyajima N.,  
 RA Tabata S.:  
 RT "Structural analysis of Arabidopsis thaliana chromosome V. VI.  
 RT Physically assigned pl and TAC clones".  
 RI DNA Res. 5:203-216(1998).  
 RI EMBL: AB012241; FAFB3044.1;  
 CR SEQUENCE: 74 AAC 4447 MW: 1364049700MBP: 499974.

Query Match	24.1%	Score 482	PR 10	Length 374
Best Local Similarity	34.0%	Pred. NO. 7.50-32		
Matches 117	Conserved 100	63	Mismatches 104	Tobias 44
				Days 12

[illegible]

```

RESULT 15
QFNLN8
ID QFNLN8 PRELIMINARY; PRI; 386 AA.
AC QFNLN8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, last annotation update)
DE S-MENOSYL-L-METHIONINE-SAM-INITIO ACIP CAPNOXYL METHYLTRANSFERASE-LIKE
DE PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
CC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
CC Magnoliopsida; eudicotyledons; core eudicots; Rosidae; eurosids 11;
CC Brassicales; Brassicaceae; Arabidopsis.
CC NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=94296549; PubMed=9428592;
RA Sato S., Kaneko T., Kotani H., Nakamura Y., Asamizu E., Miyajima N.,
RA Tabata S.;
RI "Structural analysis of Arabidopsis thaliana chromosome 5. IV.
RI Sequence features of the regions on 1456,315 bp covered by nineteen

```

RT physically assigned P1 and TAC clones.  
PL LNA Res. 5:41-54(1998).  
DR EMBL; AB010071; BAB08594.1; -  
SQ 5'GCTTTC 38; AA: 42054 MW: 100000256

Quarry Match	26-08; Score 481; TR 10; Length 386;
Best local similarity	32-8%; Fred. No. 9.4e-32;
Matches 125; Conservative	56; Mismatches 134; Indels 66; Caps 12

[illegible]

Search completed: August 4, 2004, 15:04:50  
Job time: 378 sec

Sat Sep 1 11:45:37 2001

us-09-577-657a-1.rsp